TITLE: "Social network fragmentation and community health"

AUTHORS
Goylette F. Chami[1,2]*, Sebastian E. Ahnert[3,4], Narcis B. Kabatereine[5,6], Edridah M. Tukahebwa[6]

AFFILIATIONS
[1] Dpt. Land Economy, University of Cambridge, Cambridge CB3 9EP, UK
[2] Dpt. Pathology, University of Cambridge, Cambridge CB2 1QP, UK
[3] Theory of Condensed Matter, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE UK
[4] Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR, UK
[5] Schistosomiasis Control Initiative, Imperial College London, London, W2 1PG, UK
[6] Uganda Ministry of Health, Vector Control Division, Bilharzia and Worm Control Programme, Kampala, UG

*Corresponding author.
Dpt. Land Economy
19 Silver St.
Cambridge, U.K. CB3 9EP
01223768064
gjc36@cam.ac.uk

CLASSIFICATIONS: Physical sciences; biological sciences

ABSTRACT

Community health interventions often seek to intentionally destroy paths between individuals to prevent the spread of communicable diseases. Immunizing individuals through direct vaccination or the provision of health education prevents pathogen transmission and the propagation of misinformation concerning medical treatments. Yet, it remains an open question whether network-based strategies should be used in place of conventional field approaches to target individuals for medical treatment in low-income countries. We collected complete friendship and health advice networks in 17 rural villages of Mayuge District, Uganda. Here we show that acquaintance algorithms, i.e. selecting neighbors of randomly selected nodes, were systematically more efficient in fragmenting all networks than targeting well-established community roles, i.e. health workers, village government members, and schoolteachers. Additionally, community roles were not good proxy indicators of physical proximity to other households or connections to many sick people. We also show that acquaintance algorithms were effective in offsetting potential noncompliance with deworming treatments for 16,357 individuals during mass drug administration (MDA). Health advice networks were destroyed more easily than friendship networks. Only an average of 32% of nodes were removed from health advice networks to reduce the percentage of nodes at risk of refusing treatment in MDA to below 25%. Treatment compliance of at least 75% is needed in MDA to control human morbidity attributable to parasitic worms and progress towards elimination. Our findings point toward the potential use of network-based approaches as an alternative to role-based strategies for targeting individuals in rural health interventions.

KEY WORDS: Social networks; complex networks; fragmentation; percolation; mass drug administration; health; immunization

SIGNIFICANCE STATEMENT: Fragmentation of social networks is needed in large-scale treatment campaigns. Direct vaccination of key individuals or the strategic provision of health education can prevent, respectively, the spread of viruses or misinformation. We present an easily implementable and generalizable network-based strategy for targeting households to induce fragmentation in social networks of low-income countries. Complete friendship and health advice networks were collected from 17 rural villages in Uganda. We discovered that acquaintance algorithms outperformed conventional field-based approaches for inducing social network fragmentation. Acquaintance algorithms targeted the neighbors of randomly selected nodes, whereas the latter method concerns targeting well-established community roles such as lay health workers, village government leaders, and schoolteachers. This algorithm also was effective in offsetting potential noncompliance to deworming treatments.

/body

The functioning of any complex system relies on its ability to respond to perturbations or failures(1-5). Whether high or low error tolerance of a complex system is desirable for public goods is dependent on the type of system studied. For example, high error tolerance is needed in Internet routing networks(1) and ecosystems(5), respectively, to protect against virus attacks and to ensure ecological stability after species loss. On the other hand, immunization campaigns seek to intentionally cause failures in social networks, i.e. stopping diffusion, by vaccinating a subset of individuals to quell the transmission of infectious diseases(4, 6-9).

Targeting nodes to induce fragmentation has thus far relied on the availability of information concerning network structure. The random removal of nodes requires no topological information and is a poor strategy for fragmenting complex networks(1). In contrast, targeted attacks are most detrimental to network connectivity(1, 2, 6, 7, 9). The removal of a small percentage of nodes in order of degree(1, 7, 9) or betweenness(10), particularly recalculated degree or betweenness(2), substantially damages complex networks. Although efficient, targeted attacks require full information about the global network structure. Complete network information usually is unavailable to policymakers(11). Network data can be costly to obtain, dependent on recollecting data as networks are dynamic and change over time, reliant on network type, contingent on available expertise to analyze graphs, and impractical to retrieve for time-constrained health interventions in low-income settings.

Efficient and practical network fragmentation has been achieved with acquaintance strategies(12). These algorithms target the neighbors of randomly selected nodes and utilize limited, local network information(12-14). Yet, it remains an open question whether acquaintance strategies should be employed to fragment social networks in low-income settings. Acquaintance strategies need to be compared against conventional field-based approaches. In practice, individuals are targeted in rural villages using community roles(15) as opposed to network position(16). For example, lay health workers, local government members, and schoolteachers are provided health education to stop the spread of rumors or to address concerns during en masse deworming programmes implemented in over 70 countries(17-19).

We compare the efficiency of acquaintance strategies to the efficiency of targeting community roles for damaging social networks in rural Uganda. To our knowledge, this study is the first to examine fragmentation of social networks from a low-income country. There are studies of network diffusion(15) in low-income countries. Yet, whom to target to reach the most people differs from whom to target to efficiently destroy a social network(16, 20). Importantly, diffusion based approaches that utilize a small set of seed nodes often do not reach everyone in the community(15, 21).

Two types of social networks were measured. Complete friendship and health advice networks were collected for nearly all households (3,491) in 17 villages bordering Lake Victoria in Mayuge District, Uganda. Undirected networks were graphed between households. Nodes were removed using two sets of algorithms as shown in Figure 1. All strategies first selected a node randomly. The acquaintance strategies then entailed removing either a random neighbor or a neighbor with degree ≥2 of the randomly selected node. For the acquaintance-degree strategies, a neighbor of higher degree or the highest-degree neighbor was removed. By contrast, the formal position strategy directly targeted households in order of village position; the order was first current health workers then village government members and lastly schoolteachers. When no formal positions remained, an acquaintance or acquaintance-degree strategy was employed. Efficiency was defined as the percentage of nodes required for achieving a specified level of damage to the network. Fragmentation was measured using the normalized Borgatti *F*(20, 22) indicator to capture the number and size of components remaining in the network. A completely undamaged network had F=0, whereas a destroyed network had F=1(22). Importantly, we tested the algorithms with data from a round of mass drug administration (MDA) for intestinal schistosomiasis and hookworm that we tracked at the time the social networks were surveyed(21). The percentage of nodes at risk of receiving misinformation about bad drug side effects was examined. We identified connected components that included a household with someone who refused deworming treatment due to previously experiencing an adverse drug reaction. The total number of nodes in a component with a non-compliant household was divided by the total number of nodes in the original network. Here we show that acquaintance-degree algorithms were systematically more efficient than targeting well-established community roles for fragmentation. Acquaintance-degree algorithms also were effective in reducing the percentage of nodes at risk of noncompliance in MDA.

RESULTS

Acquaintance strategies outperform targeting formal positions

Figures 2-3 present the fragmentation outcomes for a sample of four villages; 13 villages are presented in Figures S1-S2. The removal of the number of nodes that equaled the number of formal positions was examined (Table S1). Only 8-26 nodes were removed per village, which was an average of 8.70% (std. dev. 3.31%) and 8.96% (std. dev. 3.47%) of the total nodes in friendship and health advice networks, respectively. Acquaintance-degree strategies, e.g. selecting the highest degree neighbor of a random node, achieved greater fragmentation than the formal position strategy in 94.12% of all friendship (16/17) and health advice (16/17) networks. In all 17 friendship networks, removal of highest degree neighbors (F=0.204, std. dev. 0.083) induced more damage than the formal position strategy (F=0.185, std. dev. 0.067, paired t-statistic 3.395, p-value=0.004). This difference was stark for health advice networks. An average 0.412 (F std. dev. 0.207) fragmentation was achieved by removing highest degree neighbors compared to an average 0.300 (*F* std. dev. 0.136) damage caused from targeting households with formal positions (Obs. 17, paired t-statistic 5.303, p-value<0.001).

Notably, pure acquaintance strategies outperformed targeting community roles. Removing random neighbors of randomly selected nodes caused more fragmentation than targeting formal positions in 88.24% (15/17) and 76.47% (13/17), respectively, of friendship and health advice networks. There was no discernible difference between formal position types with respect to the damage caused to friendship networks; targeting any position linearly increased fragmentation. Surprisingly, removing lay health workers only caused a nonlinear change in fragmentation in 23.53% (4/17) of health advice networks.

All acquaintance and formal position strategies are attempts to heuristically approximate targeted attacks by degree(7, 9). Targeted attacks by degree, as widely shown elsewhere(1, 2), were more efficient than any acquaintance or formal position strategy for destroying all networks (Figure S3). Formal positions were a good proxy indicator for nodes with high degree; individuals with community roles as health workers, government members, and schoolteachers had on average higher degree than other households in the same village (Table S2). However, acquaintance-degree strategies more reliably selected higher degree nodes in all 34 networks than the formal position strategy (Figure 4; Figure S4). This result is remarkable considering that formal positions targeted intuitively predefined network hubs, e.g. health workers in health advice networks or community-elected village government leaders in friendship networks. Moreover, in 76.47% (13/17) of friendship networks, every node selected by the highest degree neighbor algorithm had degree greater than or equal to any node removed with a community role. In 58.82% (10/17) of health advice networks, the degree of the first node selected with the highest degree neighbor strategy was greater than or equal to the degree of the most connected node with a formal position (here, health workers).

These results suggest that limited topological information may be used in place of sociodemographic data when selecting households to reduce social network connectivity.

Physical proximity and connectivity to sick people

A role-based strategy may be employed during disease outbreaks as a potential proxy indicator of physical proximity and connections to sick people—two factors pertinent to the spread of pathogens that are directly transmitted from human-to-human. We measured how well the formal position strategy approximated physical proximity and connectivity to sick people. Physical proximity was measured as the average distance in meters from the household of interest to every other household in the village. The average distance between any two households in each village was small; households were only separated by an average of 130.61-638.67 meters (std. dev. 86.02-450.41). This result suggests that physical distance is unlikely to be a barrier to personal contact within a village. In 94.12% (16/17) of villages, households with formal positions were not significantly (p-value>0.05) closer in physical proximity to other households when compared to the average physical proximity of all households without formal positions (Table S3). We also compared the physical proximity of households with formal positions to the households selected through acquaintance-based strategies or simple random selection. Neither acquaintance-based nor formal position strategies consistently selected households with close physical proximity when compared to the selection of households at random (Figure S5). None of the proposed fragmentation strategies, including targeting formal positions, selected households that were good indicators of physical proximity to a large number of households or people.

We collected data on the number of all individuals within each home that were reported by the household head and/or wife to have diarrhea within the three months preceding the sociometric survey. In our 17 study villages, 12.44% (2035/16357) of individuals reported diarrhea. We calculated the number of people with diarrhea in the neighborhood of a node (in households directly connected to the node of interest) and divided this number by the degree of the node of interest, henceforth referred to as sickness connectivity. Diarrheal cases are of interest here because of the direct transmission of pathogens from human-to-human via the fecal-oral route, the low coverage of improved sanitation amongst all study households (12.58% (439/3491); see https://www.wssinfo.org for definition of improved sanitation), and recent large-scale cholera outbreaks in the study area(23). If physical proximity is irrelevant for contact within the study villages, we might assume that close friendships and whom individuals turn to when they are sick are proxy indicators for village contact. Accordingly, we examined how well acquaintance-based versus formal position strategies selected households with high sickness connectivity in the friendship and health advice networks. We found that an acquaintance-based strategy, here the highest-degree neighbor of a random node, more often selected households with higher sickness connectivity in friendship and health advice networks than targeting formal positions (Figure 5; Figure S6). This result is surprising considering that community health workers had high sickness connectivity, as expected, in health advice networks.

Combined formal position and acquaintance strategies

There is a practical constraint to completely destroying network connectivity (F~1) by targeting formal positions. As previously discussed, only a small proportion of households in each village had community roles. The maximum fragmentation achieved by targeting formal positions was F=0.608 in a health advice network (ID 6). Supplementing the removal of formal positions with acquaintance or acquaintance-degree strategies resulted in these network-based approaches becoming equivalent to the random removal of nodes (Figures 2-3, Figures S1-S2). Nearly all nodes were removed to completely fragment the networks. Selecting nodes first by formal position then using the highest degree neighbor strategy required on average 82.79% (std. dev. 5.25%) and 74.06% (std. dev. 7.08%) of nodes to be removed, respectively, before friendship and health advice networks were destroyed. Comparatively, using just the highest degree neighbor algorithm, all networks were more efficiently destroyed (Obs. 17, paired t-statistic friendship=19.515, health=17.910, p-values<0.001). Only an average of 62.33% (std. dev. 4.01%) and 50.35% (std. dev. 5.84%) of nodes in friendship and health advice networks, respectively, were selected to induce complete fragmentation. If a higher degree neighbor was not found then the initially selected random node was removed. This approach was not only more straightforward, but also more efficient than conventional methods(14) that ignore nodes that do not meet, for example, degree cutoffs (Figure S7). These results suggest that the inclusion of individuals with established community roles may be limited to only a probabilistic selection with acquaintance or acquaintance-degree algorithms as opposed to directly targeting formal positions.

The efficiency of acquaintance-based strategies depended on a widely noted phenomenon in social networks, the 'friendship paradox'(24). On average, the friends of a node are more connected than the node of interest(24). This disassortativity, i.e. negative degree-degree correlations, can exist because of structural constraints(25). Hubs have many connections and the number of edges that are possible between hubs is limited. As most nodes have low degree in real-world networks(26), hubs are the acquaintances of many poorly connected nodes. Thus, with a uniform probability of selection, initially nodes with a few connections will be chosen. It is then likely a hub will be sampled from the neighborhood of this peripheral node. Targeting nodes with formal positions removed significant (p-value<0.05) negative correlations of degree with average neighbor connectivity in 58.82% (10/17) of friendship and 94.12% (16/17) of health advice networks (Figures S8-S9).

Resilience by social network type

All networks displayed heavy-tailed degree distributions when compared to random networks with the same number of nodes and edges (Table S4). Though, friendship networks were more resilient than health advice networks. Within every village, removing the same percentage of nodes achieved less fragmentation in the friendship network than the health advice network (rightward shift of Borgatti F curves, Figures 2-3; Figures S1-S2). The maximum number of connected components produced by fragmenting friendship networks (avg. 12.674, std. dev. 5.022) also was less than the number (avg. 15.484, std. dev. 6.419) observed in health advice networks (Paired t-statistic -5.544, p-value<0.001).

Differences in fragmentation were due to variations in global network topology. Friendship networks (avg. N= 202.118, std. dev. 85.761) were slightly larger than health advice networks (avg. N= 193.059, std. dev. 83.069, paired t-statistic 5.306, p-value<0.001). Network transitivity, i.e. global clustering, was greater in friendship (avg. 0.113, std. dev. 0.050) than in health advice networks (avg. 0.076, std. dev. 0.033, paired t-statistic 4.834, p-value<0.001). Targeting formal positions caused little damage to friendship networks because of their almost onion-like structure(27). The core numbers of nodes with formal positions were largest (p-value<0.05) in friendship networks for 94.12% (16/17) of villages (Table S5). Hence, nodes with formal positions belonged to densely connected nuclei(16, 28) of friendship networks that remained connected after the removal of a few nodes. Acquaintance-based strategies also initially produced little damage to friendship networks due to the existence of paths between nodes of the same or lower degree. Such paths form layers around the core of the network and are onion-like in that each degree layer must be targeted(27). The removal of the core leaves the network relatively intact, as many nodes do not rely on paths through hubs to remain connected. Friendship networks (avg. 0.491, std. dev. 0.102) had a greater index(29) of onion likeness than health advice networks (avg. 0.240, std. dev. 0.099, paired t-statistic 8.565, p-value<0.001).

Fragmentation efficiency for mass drug administration

The effect of fragmentation on health outcomes is shown in Figures 6-7 and Figures S10-S11. There were 1-28 households per village (14/17) with individuals who refused deworming treatment due to previously experiencing an adverse drug reaction (Table S6). We examined the percentage of nodes in a connected component with a household that refused deworming treatment, herein referred to as being 'at risk'. This outcome was similar to the standard percolation measure of the percentage of nodes remaining in the largest component (Figure S12). Nodes were at risk because they were reachable for the spread of misinformation or negative health influences from non-compliant households(30). Hence, we investigated how the strategic ‘removal’ of households, for example by potentially providing health education before MDA, may prevent the flow of information from non-compliant households to other households.

When the same number of nodes was removed as there were formal positions, acquaintance-degree algorithms outperformed the formal position strategy in 78.57% (11/14) of friendship and 100% (14/14) of health advice networks with non-compliant households. Though, in friendship networks, only a nominal difference in the percentage of nodes remaining at risk was found when the highest degree neighbor algorithm (avg. 89.68%, std. dev. 4.88%) and the formal position strategy (avg. 90.61%, std. dev. 4.40%) were compared (Obs. 14, paired t-statistic -2.572, p-value=0.023). In health advice networks, selecting highest degree neighbors of random nodes left only an average of 77.46% (std. dev. 16.76%) of nodes at risk compared to an average of 85.12% (std. dev. 7.85%) of nodes remaining at risk after targeting formal positions (Obs. 14, paired t-statistic -2.725, p-value=0.017).

Treatment compliance of 75% is needed in MDA to control human morbidity attributable to parasitic worms and progress towards elimination(31). Accordingly, we examined what percentage of nodes must be removed for 25% or less of all nodes to remain at risk of refusing treatment. With the highest degree neighbor algorithm, only an average of 47.30% (std. dev. 5.37%) of nodes in friendship networks needed to be removed for 25% or less of nodes to be at risk. In contrast, an average 61.91% (std. dev. 17.18%) of nodes in friendship networks were removed using a combined formal position and highest degree neighbor strategy (Obs. 14, paired t-statistic -3.351, p-value=0.005). Remarkably, by selecting highest degree neighbors in health advice networks, only an average of 32.08% (std. dev. 9.48%) of nodes were removed to reduce the percentage of nodes at risk to 25%. In comparison, an average 54.54% (std. dev. 14.08%) of nodes had to be targeted by formal positions (Obs. 14, paired t-statistic -6.606, p-value<0.001). Thus, by selecting only 32%-47% of households with the highest degree neighbor algorithm in health and friendship networks, respectively, an additional 28-43% of households might be deterred from refusing treatment despite receiving no direct public intervention.

DISCUSSION

We discovered that local network strategies outperformed conventional field-based approaches in damaging rural friendship and health advice networks in Uganda. Performance was measured not only in terms of general fragmentation efficiency, but also with respect to a community health intervention required for over 1.9 billion people worldwide, i.e. MDA(32). The latter outcome concerned how best to isolate nodes from households with members that refuse medicines in order to limit the reach of noncompliance with deworming treatments(31). In 17 villages, the selection of highest degree neighbors of randomly selected nodes damaged social networks more than the targeting of households with established community roles. In practice, implementation costs(33) limit the number of households that can be approached in a village; here we showed that, even with a few nodes, more fragmentation was achieved using network-based strategies than targeting formal positions. Moreover, combining acquaintance-based algorithms with targeting formal positions resulted in a loss of fragmentation efficiency and consistency. To achieve the same outcome, more nodes were removed with combined strategies than with only acquaintance algorithms. With combined approaches, acquaintance algorithms also became inconsistently ordered, degenerating to efficiencies equivalent to the random selection of nodes. This finding is striking as it indicates that important village positions, in contrast to published literature(4, 15, 16, 30, 34), might be best left untargeted for any interventions seeking to stop the spread of information, behaviours, or pathogens through a rural social network.

Acquaintance-degree algorithms are easily implementable in low-income settings. Local network information, including neighbor degree(35), can be elicited through a simple survey prompt. The number of nodes that can be selected and, in turn, the fragmentation that can be achieved with acquaintance algorithms is not constrained by the number of households with community roles. There is frequent turnover of individuals with community roles(36). Formal position approaches require the recollecting of sociodemographic data during each intervention to accurately target current village health workers, government leaders, and schoolteachers. In contrast, acquaintance algorithms(12-14) do not need to be reformulated with changes in the community or network structure over time.

Our results are the product of one study in one geographical location. Additional research is needed to replicate our results in other low-income countries. Though beyond the scope of this study, we encourage future research to calibrate our findings for disease-specific transmission models. Here we assume that the network structure is an accurate description over which transmission of information, behaviors, or pathogens occurs. Our data lends support to this assumption. Physical distance is an unlikely barrier to transmission within a village because of the short average distance in meters between any two households. We approximated direct contact between village members by measuring close friendships and whom individuals approach when they are sick. Concerning pathogen transmission, acquaintance strategies selected households that were highly connected to other households with potentially contagious individuals. These individuals reported diarrheal illness within the past three months. We also assume transmission proceeds as a simply epidemic. This general approach provides a starting point where the probability of transmission can be calibrated to reflect the type of contagion germane to different health interventions of interest.

Our findings are promising for public interventions in rural poor settings. If our results are replicated in other contexts, health policymakers may implement the acquaintance-degree algorithms to select individuals to vaccinate or to increase drug uptake in large-scale treatment campaigns. Importantly, if these strategies were implemented in the field, special attention would need to be given to the exact nature of the sociometric item, i.e. it must conform as closely as possible to a measure that predicts dyadic transmission of the pathogen (or misinformation) of interest. Network data collection also would need to include any actors who have important transmission roles, e.g. schoolteachers, but may not be formally part of the system studied. Here, we showed that health advice networks were destroyed more easily than friendship networks. Only an average of 32% of nodes were removed to reduce the percentage of nodes at risk of refusing treatment in MDA to below 25%. Health education may be provided to a subset of individuals, who are chosen by the acquaintance-degree strategy before MDA, to strategically and preemptively prevent the spread of rumors. A number of MDA programmes are progressing towards globally or regionally eliminating infections, e.g. lymphatic filariasis and schistosomiasis(18). However, elimination efforts can be halted due to discontent with lay health workers or other individuals with village positions who are formally involved in MDA implementation(30). We showed that acquaintance-degree strategies identify alternative individuals to target for resolving negative events during MDA. Future empirical studies in low-income countries should further investigate the use of network-based approaches in place of targeting established community roles to damage social networks and, in turn, to quell the transmission of information, behaviours, or pathogens.

MATERIALS AND METHODS

Python v2.7 with the NetworkX library(37) and Stata v13.1 were used to analyze and fragment the networks. All fragmentation algorithms removed each node sequentially until only one node remained in the network and begun with only connected components (no isolates). All friendship networks began with one connected component. Health advice networks for village ID 5 and IDs 11-13 had more than one connected component. We ran 100 iterations of each algorithm on 34 undirected friendship and health advice networks. If any criteria were unmet, i.e. for the acquaintance and acquaintance-degree strategies, then the initially selected node was removed. Similarly, if the initially selected node was an isolate then that node was removed. Degree was calculated as the sum of incoming and outgoing edges with reciprocated or multi-edges treated as one edge. Detailed methods are provided in the Supplementary Information.

Ethics. This study was reviewed and approved by the Uganda National Council of Science and Technology and the Cambridge University Human Biological Research Ethics Committee. Informed consent was obtained from all respondents. Project-assigned village IDs were used to preclude the identification of individuals.

Data availability. All relevant data are available in the paper, supplementary information, and upon request from the corresponding author.

Acknowledgements. Profs. Alan Fenwick, Andreas A. Kontoleon, David W. Dunne, and Erwin Bulte were instrumental to the setup of the field project. We appreciate the involvement of the study villages in Mayuge District, Uganda. Several field teams made this project possible: the laboratory technicians from the Vector Control Division in Uganda who collected the parasitology, the field assistants, and local village surveyors who collected household data. This study was financially supported by the Vice Chancellor's Fund of the University of Cambridge, the Schistosomiasis Control Initiative, the Wellcome Trust Programme grant 083931/Z/07/Z to David W. Dunne, the Netherlands Organization for Scientific Research grant 452-04-333 to Erwin Bulte, and the Isaac Newton Trust and King's College, Cambridge Fellowships to Goylette F. Chami.

References

1 Albert, R., Jeong, H. & Barabási, A.-L. (2000) Error and attack tolerance of complex networks. *Nature* 406(6794): 378-382

2 Holme, P., Kim, B. J., Yoon, C. N. & Han, S. K. (2002) Attack vulnerability of complex networks. *Phys Rev E* 65(5): 056109

3 Fornito, A., Zalesky, A. & Breakspear, M. (2015) The connectomics of brain disorders. *Nat Rev Neurosci* 16(3):159-172

4 Scarpino, S. V., Allard, A. & Hebert-Dufresne, L. (2016) The effect of a prudent adaptive behaviour on disease transmission. *Nat Phys* 12(11), 1042-1046

5 Montoya, J. M., Pimm, S. L. & Solé, R. V. (2006) Ecological networks and their fragility. *Nature* 442(7100), 259-264

6 Pastor-Satorras, R. & Vespignani, A. (2001) Epidemic spreading in scale-free networks. *Phys Rev Lett* 86(14), 3200-3203

7 Pastor-Satorras, R. & Vespignani, A. (2002) Immunization of complex networks. *Phys Rev E* 65(3): 036104

8 Morone, F. & Makse, H. A. (2015) Influence maximization in complex networks through optimal percolation. *Nature* 524(7563): 65-68

9 Dezső Z. & Barabási A-L. (2002) Halting viruses in scale-free networks. *Phys Rev E* 65(5):055103.

10 Girvan, M. & Newman, M. E. J. (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99(12): 7821-7826

11 Kossinets, G. (2006) Effects of missing data in social networks. *Soc Net* 28(3): 247-268

12 Cohen, R., Havlin, S. & ben-Avraham, D. (2003) Efficient immunization strategies for computer networks and populations. *Phys Rev Lett* 91(24): 247901

13 Holme, P. (2004) Efficient local strategies for vaccination and network attack. *EPL* 68(6): 908

14 Gallos, L. K., Liljeros, F., Argyrakis, P., Bunde, A. & Havlin, S. (2007) Improving immunization strategies. *Phys Rev E* 75(4): 045104

15 Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2013) The Diffusion of microfinance. *Science* 341(6144)

16 Kitsak, M. *et al.* (2010) Identification of influential spreaders. *Nat Phys* 6(11): 888-893

17 Amazigo, U. V., Leak, S. G., Zoure, H. G., Njepuome, N. & Lusamba-Dikassa, P. S. (2012) Community-driven interventions can revolutionise control of neglected tropical diseases. *Trends Parasitol* 28(6)

18 World Health Organization (2012) Accelerating work to overcome the global impact of neglected tropical diseases: A roadmap for implementation.

19 Fenwick, A. *et al.* (2009) The Schistosomiasis Control Initiative (SCI): rationale, development and implementation from 2002-2008. *Parasitology* 136(13): 1719-1730

20 Borgatti, S. P. (2006) Identifying sets of key players in a social network. *Comput Math Organ Theory* 12(1): 21-34,

21 Chami, G. F. *et al.* (2016) Profiling nonrecipients of mass drug administration for schistosomiasis and hookworm infections: a comprehensive analysis of praziquantel and albendazole coverage in community-directed treatment in Uganda. *Clin Infect Dis* 62(2): 200-207

22 Chen, Y. *et al.* (2007) Percolation theory applied to measures of fragmentation in social networks. *Phys Rev E* 75(4): 046107

23 Bwire G, *et al*. (2017) Epidemiology of cholera outbreaks and socio-economic characteristics of the communities in the fishing villages of Uganda: 2011-2015. *PLOS Negl Trop Dis* 11(3):e0005407.

24 Feld, S. L. (1991) Why your friends have more friends than you do. *Am J Sociol* 96(6): 1464-1477

25 Maslov, S., Sneppen, K. & Zaliznyak, A. (2004) Detection of topological patterns in complex networks: correlation profile of the internet. *Phys A* 333(C): 529-540

26 Barabási, A.-L. & Albert, R. (1999) Emergence of scaling in random networks. *Science* 286(5439): 509-512

27 Schneider, C. M., Moreira, A. A., Andrade, J. S., Havlin, S. & Herrmann, H. J. (2011) Mitigation of malicious attacks on networks. *Proc Natl Acad Sci USA* 108(10): 3838-3841

28 Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y. & Shir, E. (2007) A model of Internet topology using k-shell decomposition. *Proc Natl Acad Sci USA* 104(27):11150-11154

29 Louzada, V. H. P., Daolio, F., Herrmann, H. J. & Tomassini, M. (2013) Smart rewiring for network robustness. *J Complex Networks* 1(2): 150-159

30 Krentel, A., Fischer, P. U. & Weil, G. J. (2013) A review of factors that influence individual compliance with mass drug administration for elimination of lymphatic filariasis. *PLoS Negl Trop Dis* 7(11): e2447

31 Shuford, K. V., Turner, H. C. & Anderson, R. M. (2016) Compliance with anthelmintic treatment in the neglected tropical diseases control programmes: a systematic review. *Parasit Vectors* 9(1): 29

32 Webster, J. P., Molyneux, D. H., Hotez, P. J. & Fenwick, A. (2014) The contribution of mass drug administration to global health: past, present and future. *Philos Trans R Soc Lond, B, Biol Sci* 369(1645)

33 Lo, N. C. *et al.* (2015) Comparison of community-wide, integrated mass drug administration strategies for schistosomiasis and soil-transmitted helminthiasis: a cost-effectiveness modelling study. *Lancet Glob. Health* 3(10): e629-e638

34 Valente, T. W. & Davis, R. L. (1999) Accelerating the Diffusion of Innovations Using Opinion Leaders. *Ann Am Acad Pol Soc Sci* 566(1): 55-67

35 Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2014) Gossip: Identifying central individuals in a social network. (National Bureau of Economic Research).

36 Emukah, E. C. *et al.* (2008) Factors affecting the attrition of community-directed distributors of ivermectin, in an onchocerciasis-control programme in the Imo and Abia states of south-eastern Nigeria. *Ann Trop Med Parasitol* 102(1): 45-51

37 Schult, D. A. & Swart, P. (2008) in Proceedings of the 7th Python in Science Conferences (SciPy). 11-16.

Figure Legends

Figure 1 Schematic illustrating network-based fragmentation algorithms. All four strategies begin with the random choice of a node (Step 1), before proceeding to Step 2, at which one of four strategies can be chosen. Strategies A and B are acquaintance strategies, which entail choosing a random neighbor or a random neighbor with degree ≥2 of the node chosen in Step 1. Strategies C and D are acquaintance-degree strategies. A random neighbor with degree greater than the original node (C) or the neighbor with highest degree (D) is chosen. A random choice is made between two nodes of equal highest degree.

Figure 2 Fragmentation outcomes for friendship networks. Four villages are shown that had the fewest, median, 75^{th} percentile, and greatest number of nodes. The remaining villages are shown in Figure S1. IDs correspond to project-assigned village IDs. N is the total number of nodes in the original network and FP is the total number of formal positions in the village. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

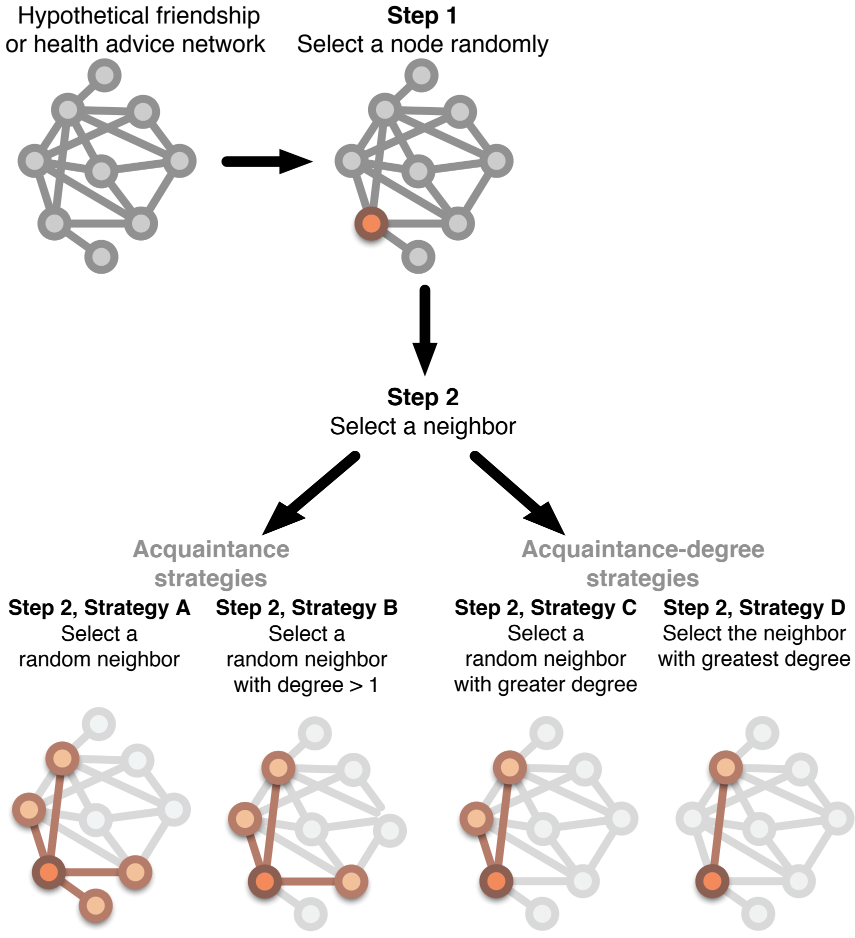
Figure 3 Fragmentation outcomes for health advice networks. Four villages are shown that had the fewest, median, 75^{th} percentile, and greatest number of nodes. The remaining villages are shown in Figure S2. IDs correspond to project-assigned village IDs. N is the total number of nodes in the original network and FP is the total number of formal positions in the village. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

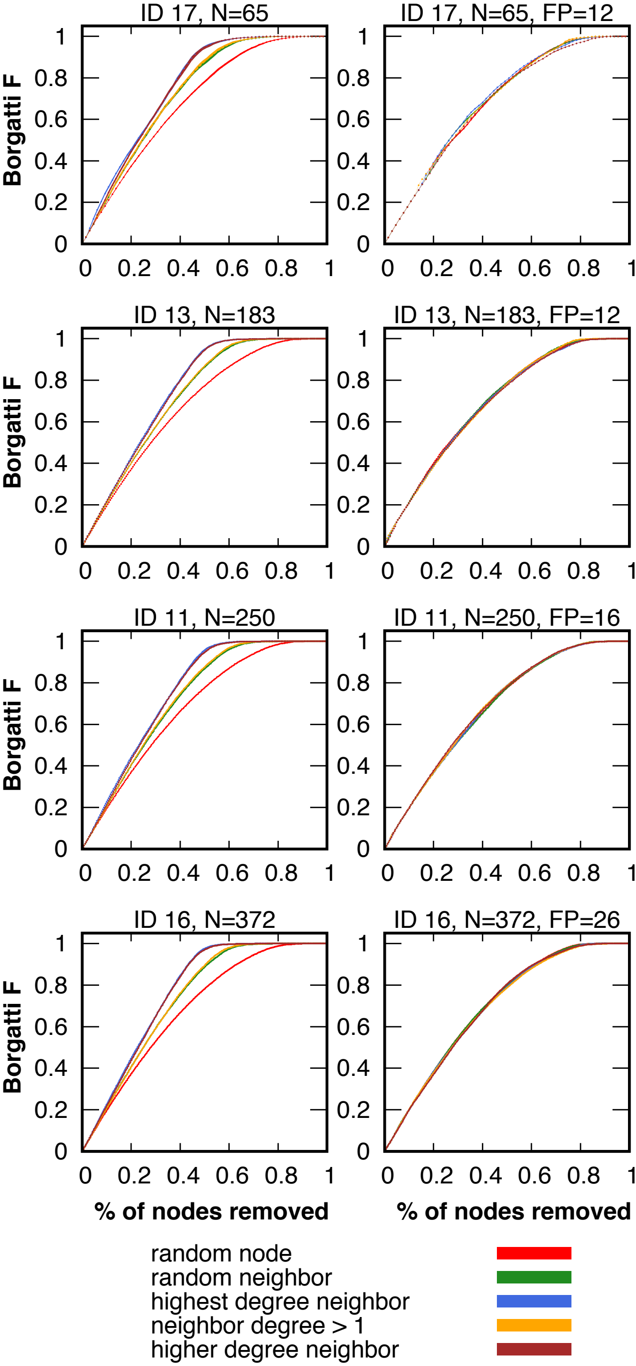
Figure 4 Avg. degree of node removed by acquaintance and formal position strategies. Four villages are shown that had the fewest, median, 75^{th} percentile, and greatest number of nodes. The remaining villages are shown in Figure S4. The average degree for each node removed is shown up to the number of formal positions. IDs correspond to project-assigned village IDs. The type of network is labeled accordingly. One thousand iterations were run and line widths represent 95% confidence intervals.

Figure 5 Avg. connectivity to sick people. Four villages are shown that had the fewest, median, 75th percentile, and greatest number of nodes. The remaining villages are shown in Figure S6. IDs correspond to project-assigned village IDs. The type of network is labeled accordingly. Sickness connectivity was defined as follows. The number of people in the neighborhood of a node who reported diarrhea within the three months preceding the sociometric survey was divided by the degree of the node of interest. The average sickness connectivity for each node removed is shown up to the number of formal positions. One thousand iterations were run and line widths represent 95% confidence intervals.

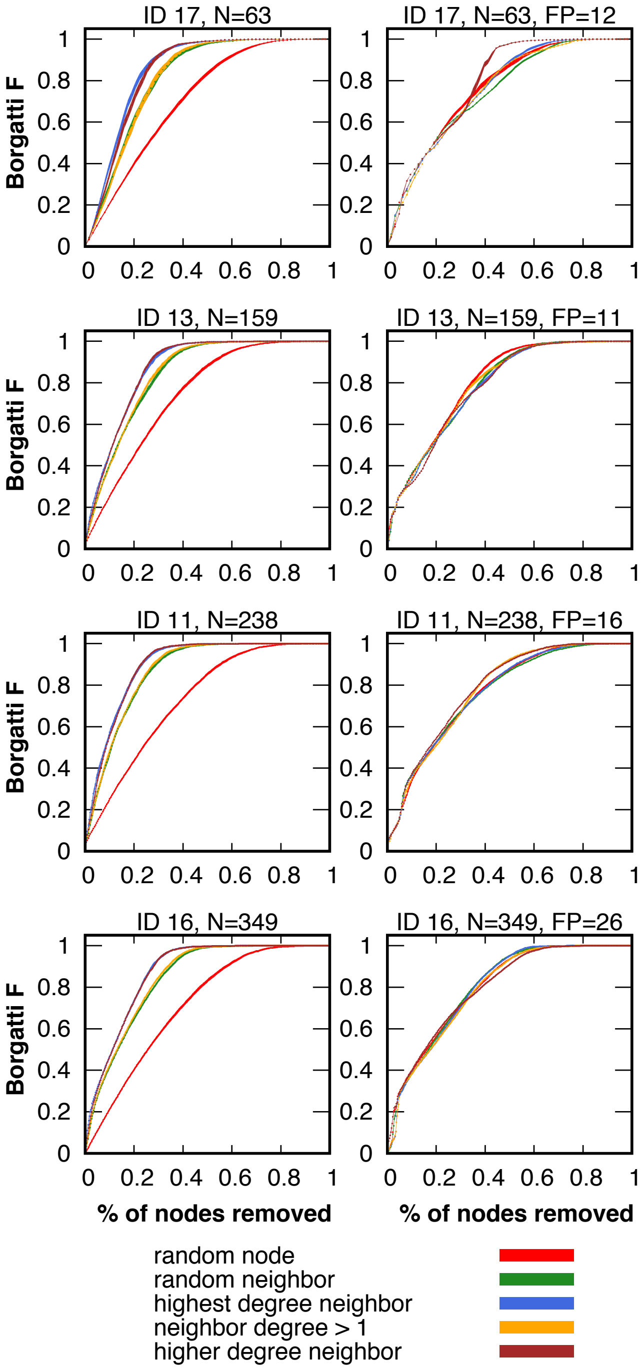
Figure 6 Health outcomes for friendship networks. Four villages are shown that had the fewest, median, 75th percentile, and greatest number of nodes. The remaining villages are shown in Figure S10. IDs correspond to project-assigned village IDs. N is the total number of nodes in the original network and FP is the total number of formal positions in the village. NS is the number of non-compliant seeds; these nodes represent households with someone who refused deworming treatment due to a previous adverse drug reaction. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

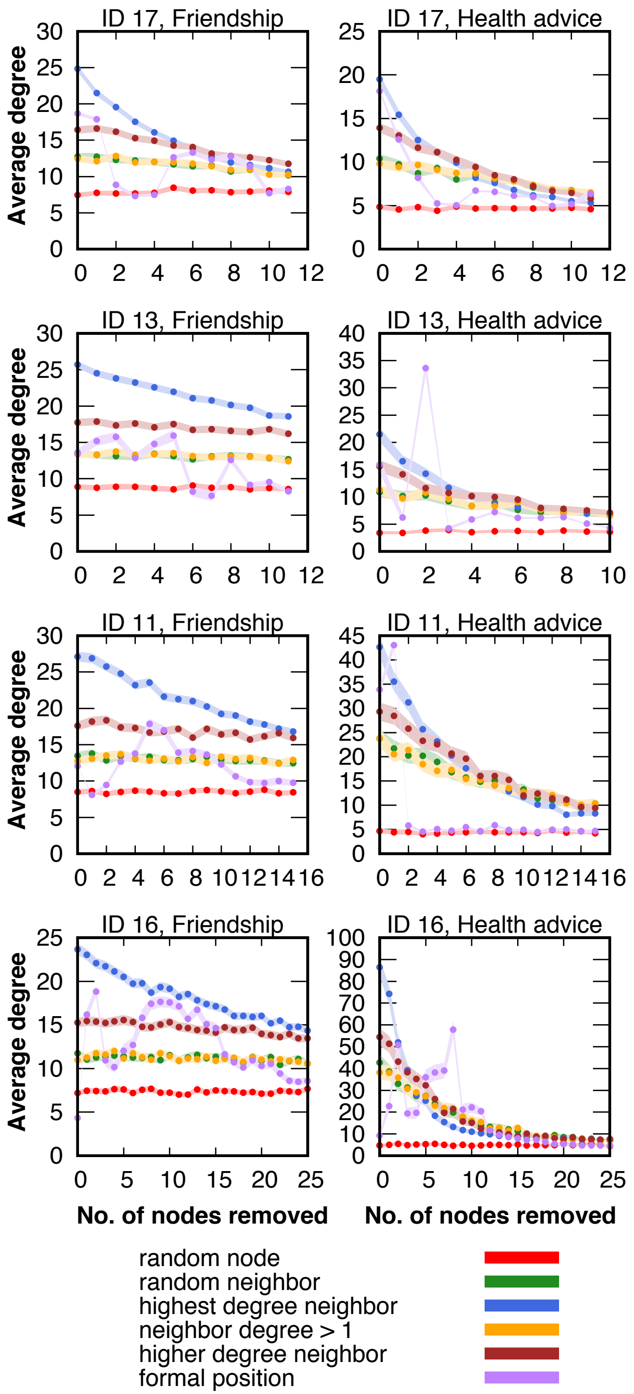
Figure 7 Health outcomes for health advice networks. Four villages are shown that had the fewest, median, 75th percentile, and greatest number of nodes. The remaining villages are shown in Figure S11. IDs correspond to project-assigned village IDs. N is the total number of nodes in the original network and FP is the total number of formal positions in the village. NS is the number of non-compliant seeds; these nodes represent households with someone who refused deworming treatment due to a previous adverse drug reaction. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

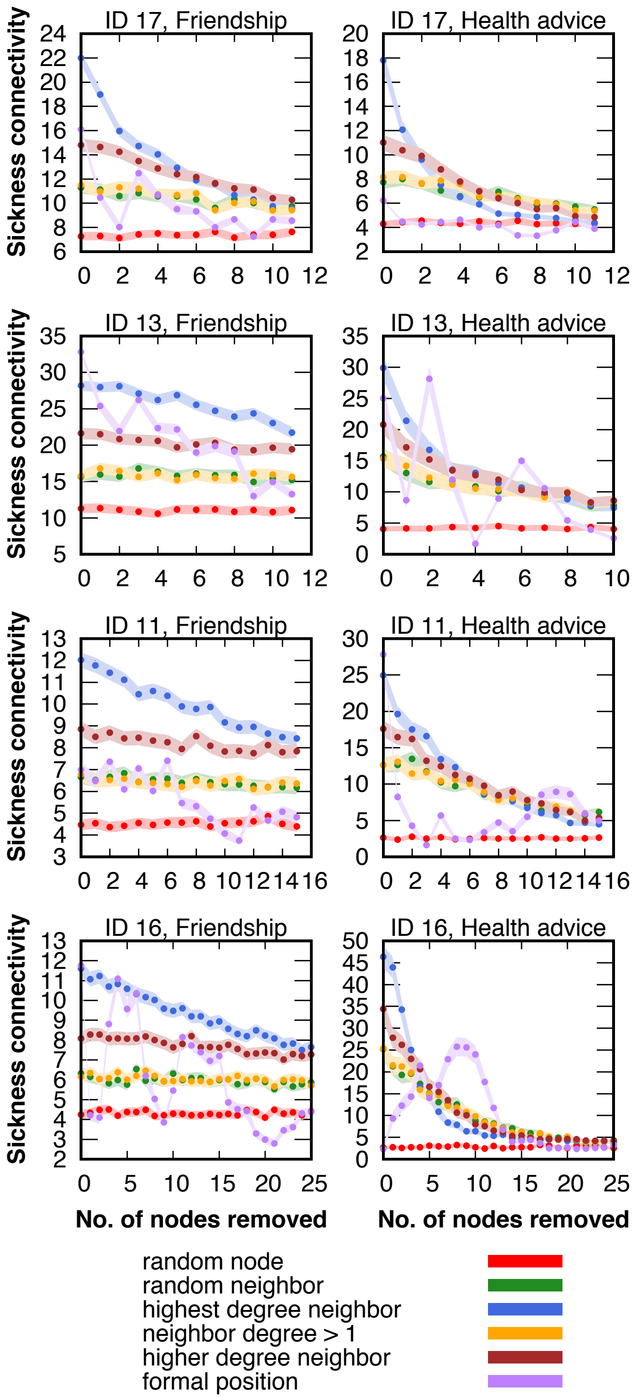

Hypothetical friendship
or health advice network
Step 1
Select a node randomly
Step 2
Select a neighbor
Acquaintance
strategies
Acquaintance-degree
strategies
Step 2, Strategy A
Select a
random neighbor
Step 2, Strategy B
Select a
random neighbor
with degree > 1
Step 2, Strategy C
Select a
random neighbor
with greater degree
Step 2, Strategy D
Select the neighbor
with greatest degree

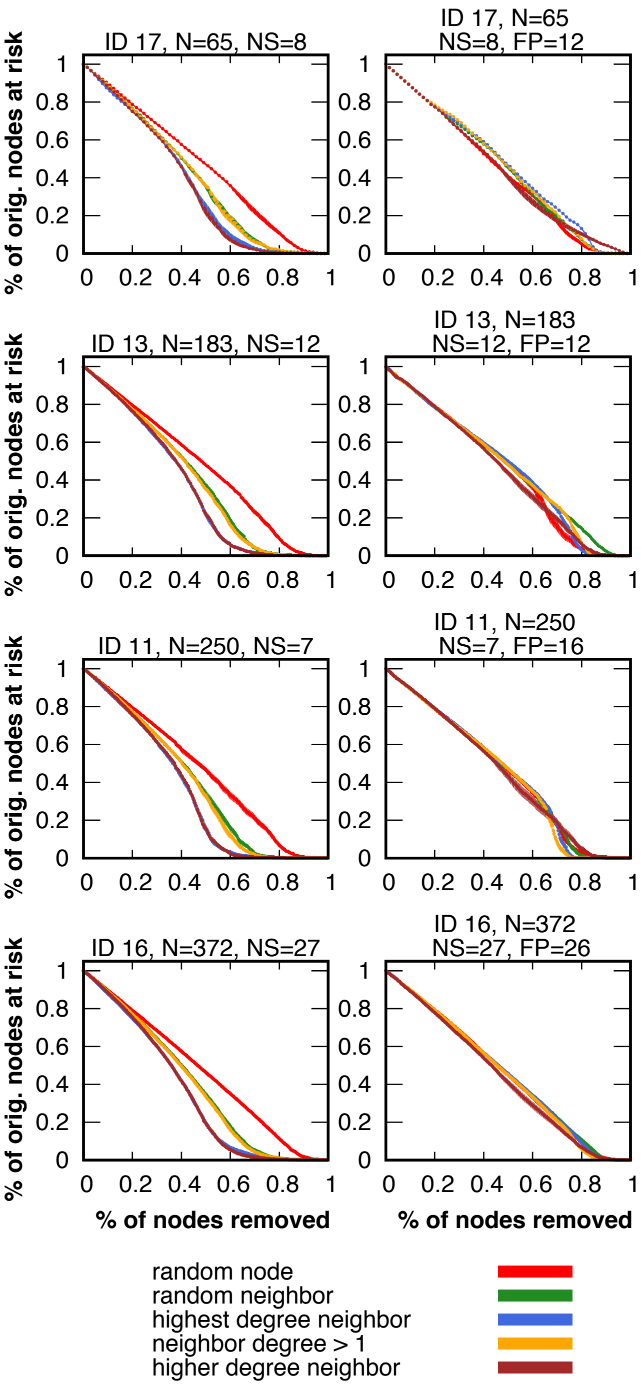


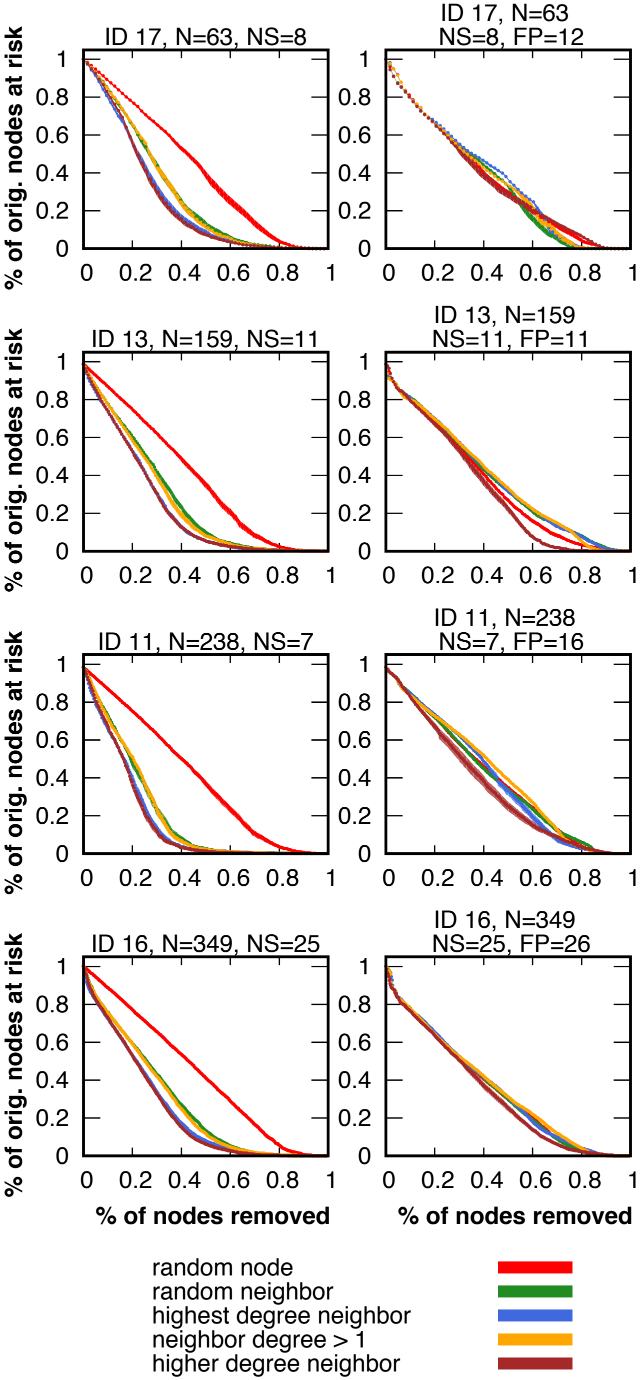
ID 17, N=65
ID 17, N=65, FP=12
ID 13, N=183
ID 13, N=183, FP=12
ID 11, N=250
ID 11, N=250, FP=16
ID 16, N=372
ID 16, N=372, FP=26
Borgatti F
% of nodes removed
0
0.2
0.4
0.6
0.8
1
random node
random neighbor
highest degree neighbor
neighbor degree > 1
higher degree neighbor


ID 17, N=63
ID 17, N=63, FP=12
ID 13, N=159
ID 13, N=159, FP=11
ID 11, N=238
ID 11, N=238, FP=16
ID 16, N=349
ID 16, N=349, FP=26
Borgatti F
% of nodes removed
random node
random neighbor
highest degree neighbor
neighbor degree > 1
higher degree neighbor


ID 17, Friendship
ID 17, Health advice
ID 13, Friendship
ID 13, Health advice
ID 11, Friendship
ID 11, Health advice
ID 16, Friendship
ID 16, Health advice
Average degree
No. of nodes removed
random node
random neighbor
highest degree neighbor
neighbor degree > 1
higher degree neighbor
formal position


ID 17, Friendship
ID 17, Health advice
ID 13, Friendship
ID 13, Health advice
ID 11, Friendship
ID 11, Health advice
ID 16, Friendship
ID 16, Health advice
Sickness connectivity
No. of nodes removed
random node
random neighbor
highest degree neighbor
neighbor degree > 1
higher degree neighbor
formal position


ID 17, N=65, NS=8
ID 17, N=65
NS=8, FP=12
ID 13, N=183, NS=12
ID 13, N=183
NS=12, FP=12
ID 11, N=250, NS=7
ID 11, N=250
NS=7, FP=16
ID 16, N=372, NS=27
ID 16, N=372
NS=27, FP=26
% of orig. nodes at risk
% of nodes removed
0
0.2
0.4
0.6
0.8
1
random node
random neighbor
highest degree neighbor
neighbor degree > 1
higher degree neighbor


ID 17, N=63, NS=8
ID 17, N=63
NS=8, FP=12
ID 13, N=159, NS=11
ID 13, N=159
NS=11, FP=11
ID 11, N=238, NS=7
ID 11, N=238
NS=7, FP=16
ID 16, N=349, NS=25
ID 16, N=349
NS=25, FP=26
% of orig. nodes at risk
% of nodes removed
0
0.2
0.4
0.6
0.8
1
random node
random neighbor
highest degree neighbor
neighbor degree > 1
higher degree neighbor

SUPPLEMENTARY MATERIAL: "Social network fragmentation and community health"

Goylette F. Chami*, Sebastian E. Ahnert, Narcis B. Kabatereine, Edridah M. Tukahebwa

Table of Contents

Materials and methods

Network prompts

Networks were generated at the household level because community medicine distributors (CMDs) were trained to and have been shown to move from door to door to deliver medicines from during MDA(1).

Close friendship: "Please tell me the clan name first then the second name of up to 10 people that are very close friends to you. You should feel comfortable to turn to this person to borrow tools for fishing or farming without paying. A close friend is also someone that you see frequently. Do not name anyone in your household. Provide the names in the order of who is your closest friend first. Only name people in your village."

Health advice: "Please tell me the clan name first then the second name of up to 10 people that you trust for advice about taking drugs or any health problems. These people do not have to be health workers. Provide the names in the order of whose opinion you value most and who you would go to first. Only name people in your village."

Fragmentation algorithms

1. Random node removal
2. Acquaintance strategy
    a. Random neighbor
    b. Random neighbor with degree ≥2
3. Acquaintance-degree strategy
    a. Highest degree neighbor
    b. Higher degree neighbor
4. Formal position strategy
    a. Random neighbor
    b. Random neighbor with degree ≥2
    c. Higher degree neighbor
    d. Highest degree neighbor

For the random node removal, all nodes had a uniform probability of selection. The acquaintance and acquaintance-degree strategies began with the selection of a random node then a neighbor (direct connection) of the initially selected node was removed. Two acquaintance algorithms were employed. Algorithm 2A randomly removed a neighbor of the initially selected node(2). In 2B, a restriction was added to 2A where the randomly removed neighbor must have a degree of at least two. This criterion is similar to setting a local threshold for the neighbor's degree(3) and guided the removal of neighbors who had a connection to at least one additional node that was not the initially selected random node. In the event that a neighbor was selected from an isolated dyad then, in 2B, this neighbor was removed. The acquaintance-degree strategy introduced a trivial improvement(3, 4) in the acquaintance strategy(2). Acquaintance-degree algorithm 3A removed the highest degree neighbor of the initially selected node(4). If there was a tie, i.e. if two neighbors had the highest degree value then one of these neighbors was randomly selected and removed. Algorithm 3B randomly removed a neighbor with higher degree than the initially selected node.

The formal position strategy purposely targeted individuals with community roles. In this strategy, we first directly removed individuals in order of village positions then, when no individuals with formal positions remained, an acquaintance or acquaintance-degree strategy was employed. Formal positions included households with at least one individual in at least one of the following categories at the time of the network survey: government health workers, CMDs who were village-elected health workers, local council members (village government), and schoolteachers. These categories reflect actual field practices in community-based MDA in Uganda(1, 5). Health personnel from outside of a village will work with influential, local stakeholders to respond to problems arising in a village during treatment campaigns. These individuals are influential because they are the implementers of community-based MDA (health workers), have high social status (local council), or are the implementers of MDA in primary schools (teachers). There was a fixed number of two CMDs per village and a maximum of nine village government members. No fixed or maximum number of government health workers or schoolteachers existed. The local council positions were as follows: chairman, vice chairman, secretary, defense, gender secretary, disabled secretary, youth council, elderly secretary, or information secretary. The ranking (hierarchy) of formal positions and order of node removal was health workers (both government and CMDs) then local council members and finally schoolteachers. Within each category of formal positions, if there were multiple individuals then one of these individuals was randomly chosen and removed. If an individual in a household held multiple formal positions or multiple individuals in a household had formal positions across different categories then the household was assigned the category with the highest ranking.

Targeted attack algorithms

1. Targeted attacks
   a. Highest degree
   b. Highest betweenness
   c. Recalculated highest degree
   d. Recalculated highest betweenness

Targeted attacks were strategies that removed nodes based on centrality(6) and required global network information. Algorithms 5A and 5B removed nodes in descending order of degree(7) and betweenness(8), respectively. The recalculated measures(9) recounted degree or updated betweennness after each node removal. For ties, i.e. the same value assigned to different nodes, a node was randomly chosen amongst nodes with the same value of degree or betweenness. Only 10 iterations were run for betweenness due to the infrequency of ties.

Fragmentation outcomes

The main outcome was the total number of fragments with adjustments for component size using the Borgatti F(10) indicator as described in Chen *et al*(11) where F=0 was an undamaged network and F=1 equaled maximum fragmentation. F asymptotically approached zero when isolates remained, so complete destruction of network connectivity was defined here as F=0.9945. A connected component was defined as a group of at least two connected nodes. To check the robustness of the acquaintance and acquaintance-degree results as well as to enable comparisons with published studies, the standard percolation outcome(3, 9, 11-13) also was calculated. The percolation outcome measured the percentage of nodes remaining in the largest component.

Health outcome
MDA is the distribution of preventive chemotherapies to an entire population within a defined geographical area and predominantly at risk of infection with one of six parasitic worms(5, 14). Over 1.9 billion individuals worldwide require treatment through MDA(14). In our study area, community-based MDA(1, 5) was used to distribute praziquantel, albendazole, and ivermectin for the treatment of intestinal schistosomiasis, soil-transmitted helminths, and lymphatic filariasis. MDA is the main, and most often only available method of controlling morbidity attributable to these infections. Yet, an adverse drug reaction experienced by a few individuals within a village can cause widespread refusal to ingest pills (noncompliance) and, in turn, destabilize or halt MDA, even at times stopping treatment for several years (15-17). Widespread noncompliance ensues ultimately from the spread of information, which can include rumours, about the adverse event(15). Considering that information travels along connections in friendship and health advice networks(18-20) and the starting points (seeds) for this diffusion are the individuals/households experiencing the adverse event then there is a need to quell the ability of these seeds to spread information to the rest of the network.

All households in the networks were interviewed to record who was offered medicine by CMDs (implementers of community-based MDA(1)) and, amongst those offered, who refused to ingest pills. Here, noncompliance included only individuals who refused to swallow medicines because of a previous experience of adverse drug side effects. A node (household) was classified as a non-compliant seed if at least one individual, who was eligible for treatment in the household, refused all pills during the MDA conducted at the time of the network survey.

We measured the percentage of nodes in the network that were at risk of receiving information from a non-compliant seed. We assume all nodes in a component with a non-compliant seed were reachable by that seed. Accordingly, we divided the total number of nodes in a connected component with a non-compliant seed by the total number of nodes in the original network.

Comparison of formal position targeting to uniform random node removal
When examining the same number of nodes removed as there were formal positions, we also compared the efficiency of formal position targeting to a simple approach that is not an acquaintance/network strategy, i.e. the uniform random sampling of households (Figures S1-S2). Targeting formal positions outperformed uniform random selection in 58.82% (10/17) of friendship and 88.24% (15/17) of health advice networks. In the friendship networks, the average fragmentation achieved with the formal position strategy (F 0.185, std. dev. 0.067) was only slightly larger than the fragmentation (F 0.180, std. dev. 0.066) observed after randomly removing households (Obs. 17, paired t-statistic 2.640, p-value=0.018). For health advice networks, targeting formal positions induced more fragmentation (avg. F 0.30, std. dev. 0.136) than that achieved with random selection (avg. F 0.196, std. dev. 0.070, Obs. 17, paired t-statistic 4.962, p-value<0.001).

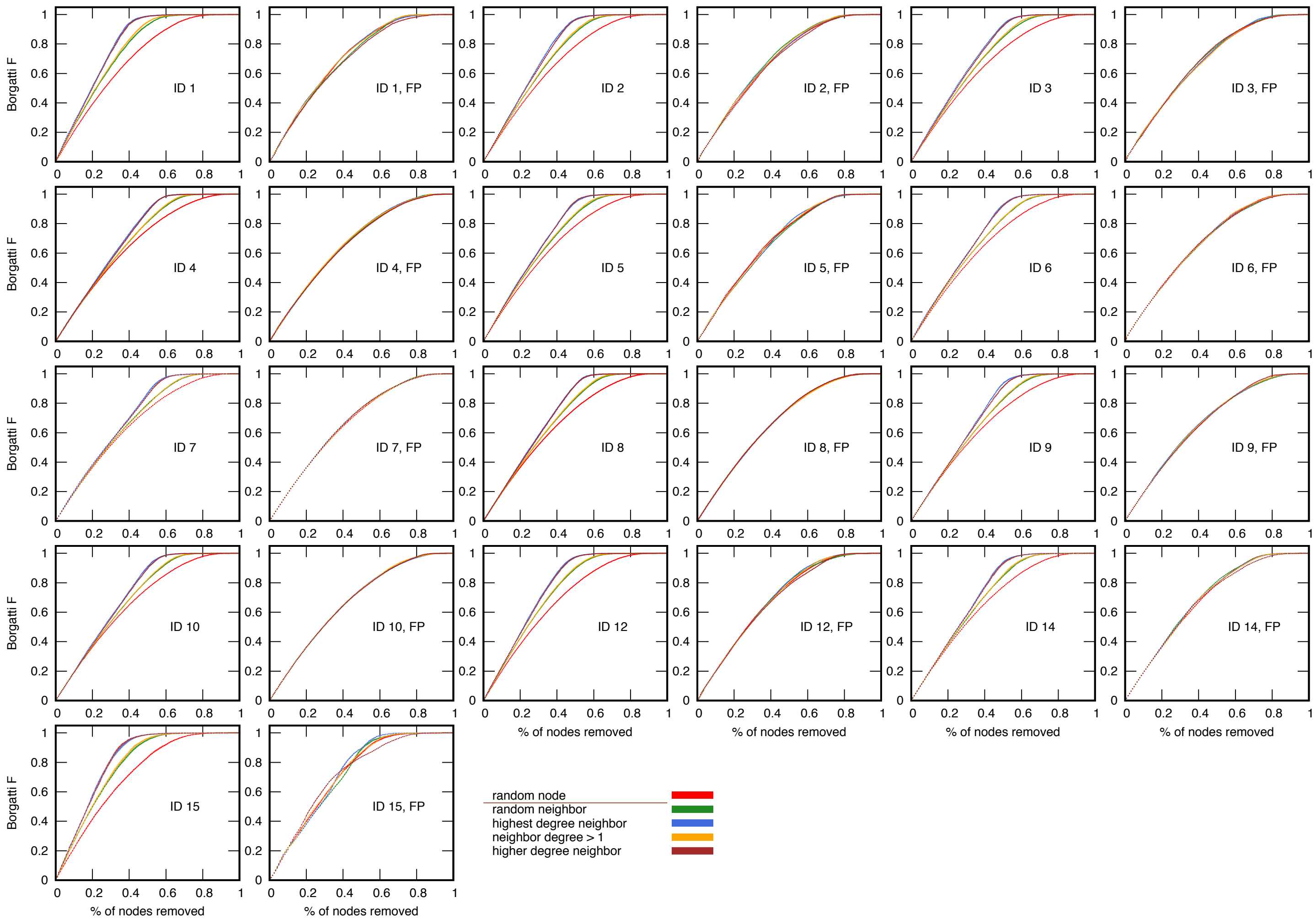


Figure S1: Fragmentation outcomes for 13 friendship networks. Thirteen villages are shown that were not presented in the main text. IDs correspond to project-assigned village IDs. N is the total number of nodes in the original network. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

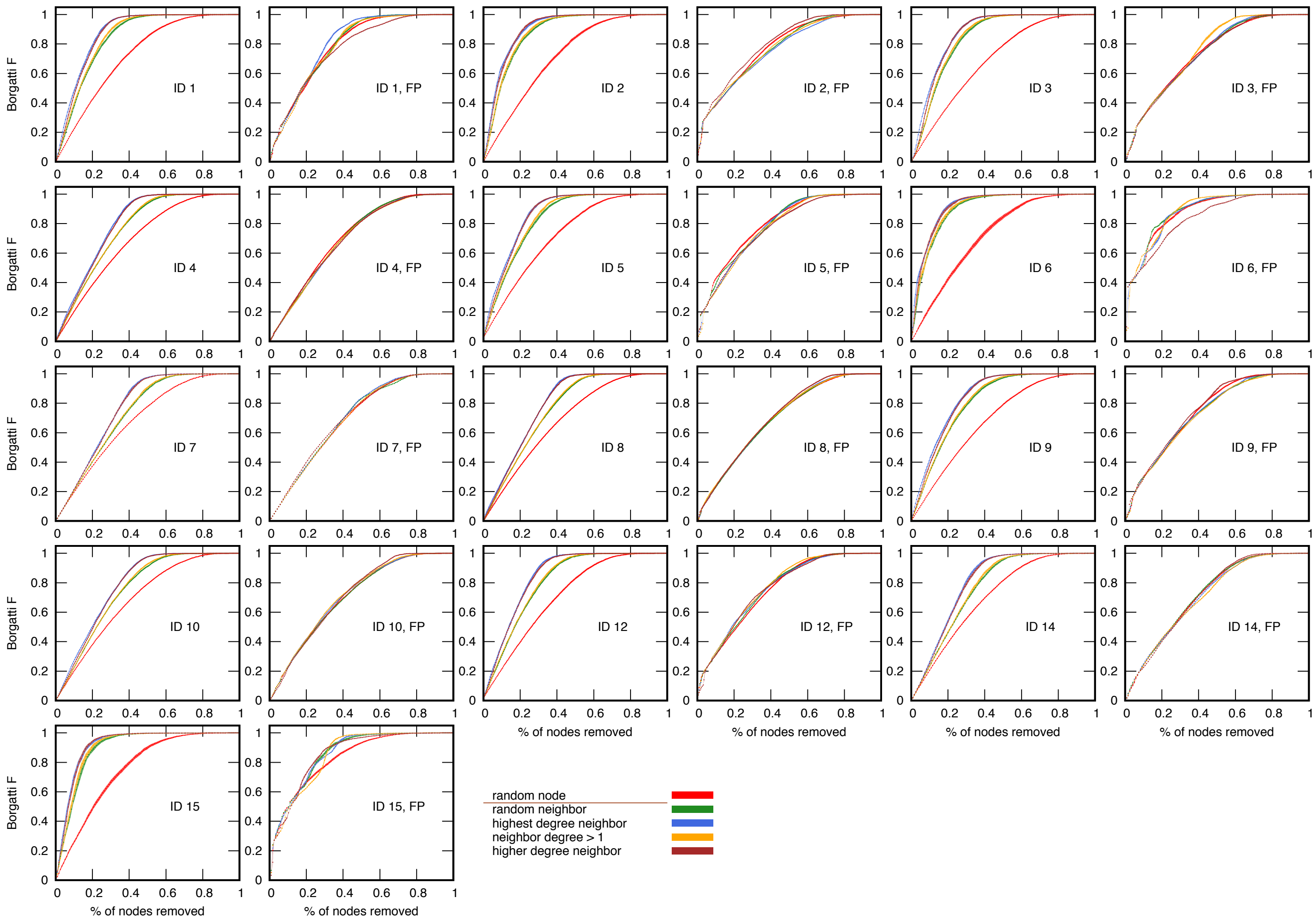


Figure S2: Fragmentation outcomes for 13 health advice networks. Thirteen villages are shown that were not presented in the main text. IDs correspond to project-assigned village IDs. N is the total number of nodes in the original network. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

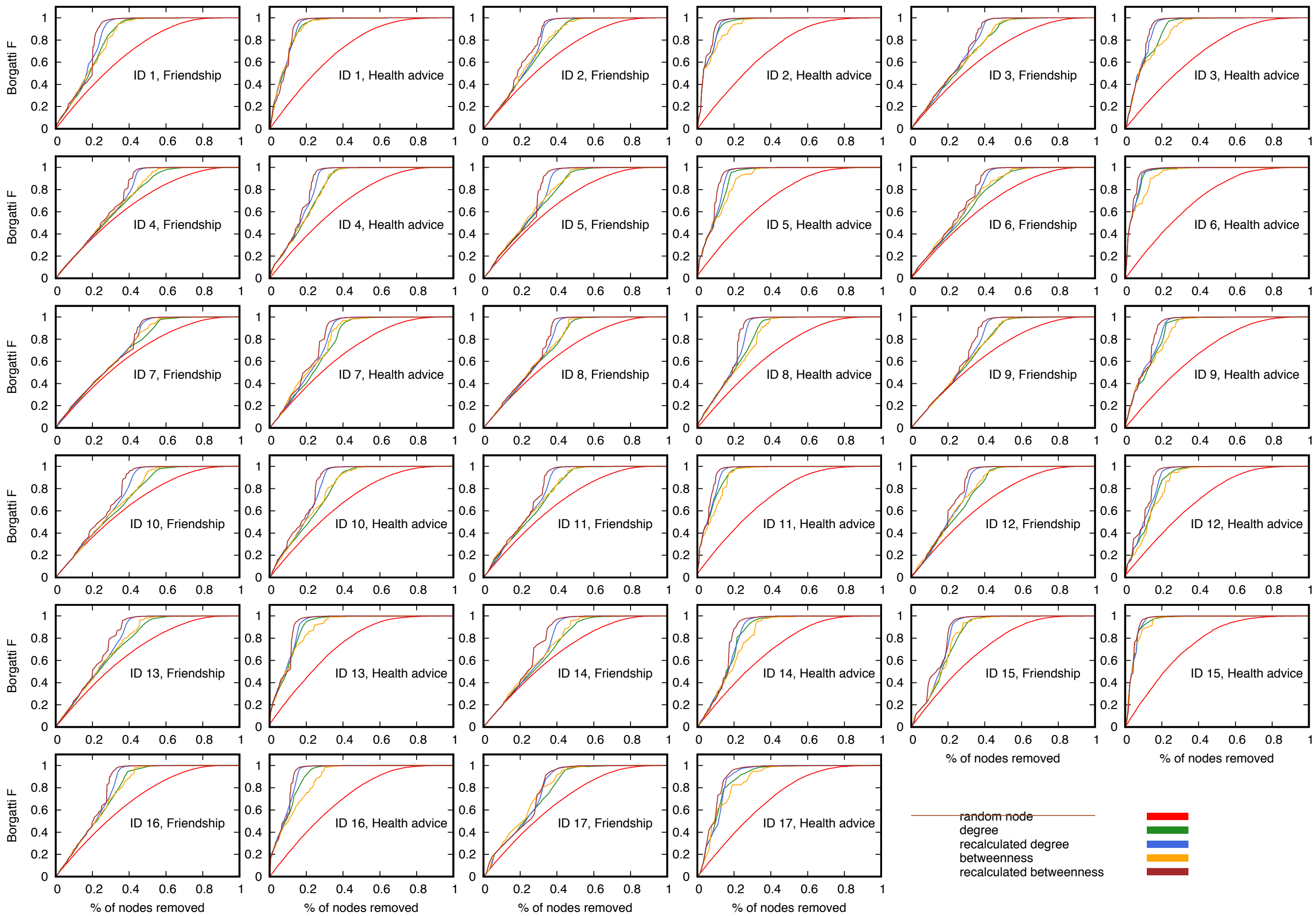


Figure S3: Targeted attacks by degree and betweenness. Fragmentation algorithms that utilized full network information are shown for all villages and networks. IDs correspond to project-assigned village IDs. Line widths are greater than the 95% confidence intervals.

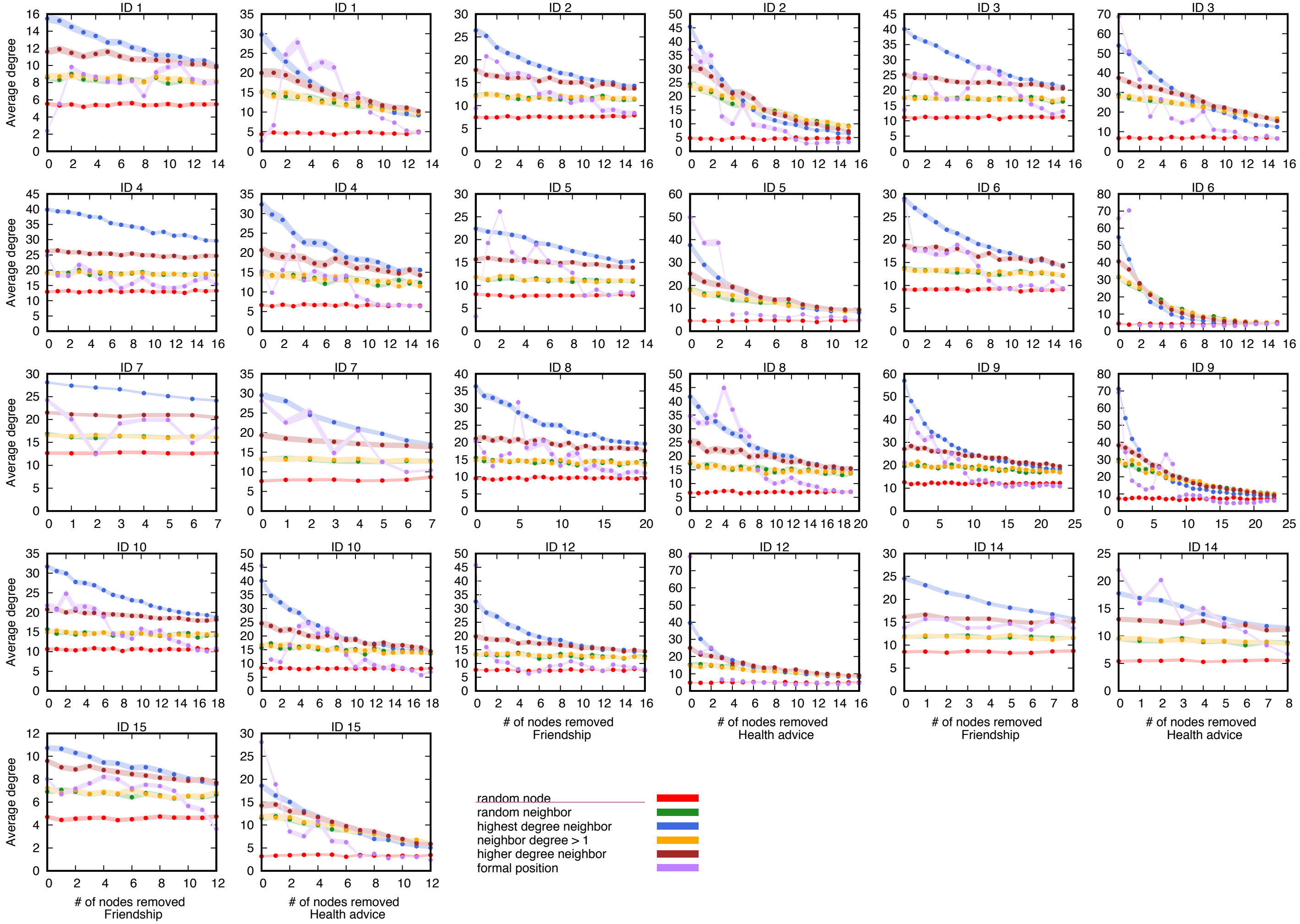


Figure S4: Avg. degree of node removed for 13 friendship and health advice networks. Thirteen villages are shown that were not presented in the main text. The average degree for each node removed is shown up to the number of formal positions. IDs correspond to project-assigned village IDs. The type of network is labeled accordingly. One thousand iterations were run and line widths represent 95% confidence intervals.

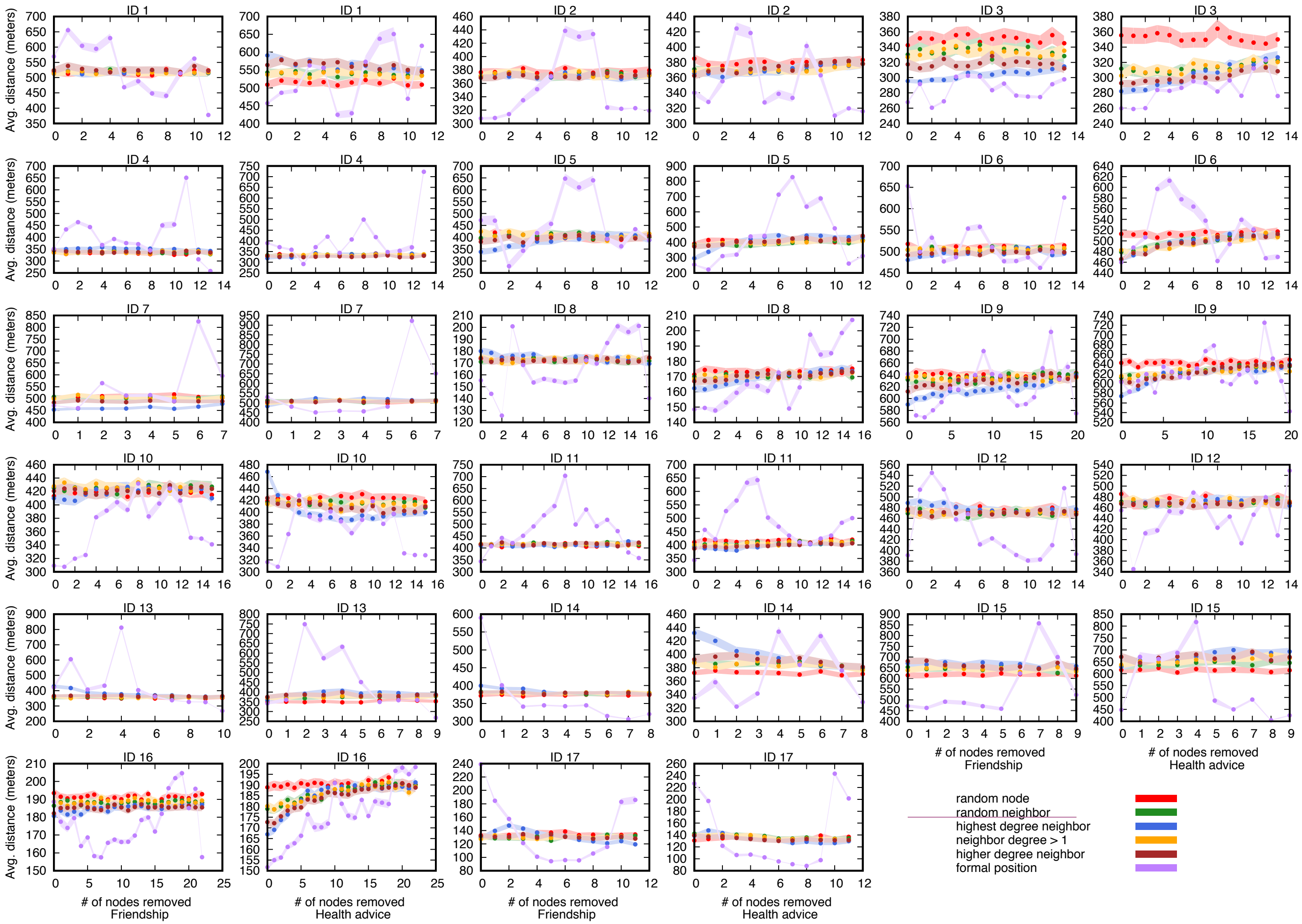


Figure S5: Physical proximity of node selected by each fragmentation strategy. The average haversine distance in meters is shown for each node selected by each fragmentation strategy. One thousand iterations were run and line widths represent 95% confidence intervals. If a neighbour was selected that did not have available GPS waypoint data then the initially selected node was removed. If both the neighbour and the initially selected node did not have available GPS waypoint data then a new initial node was selected. Only the number of nodes as there were formal positions with GPS waypoint data was removed.

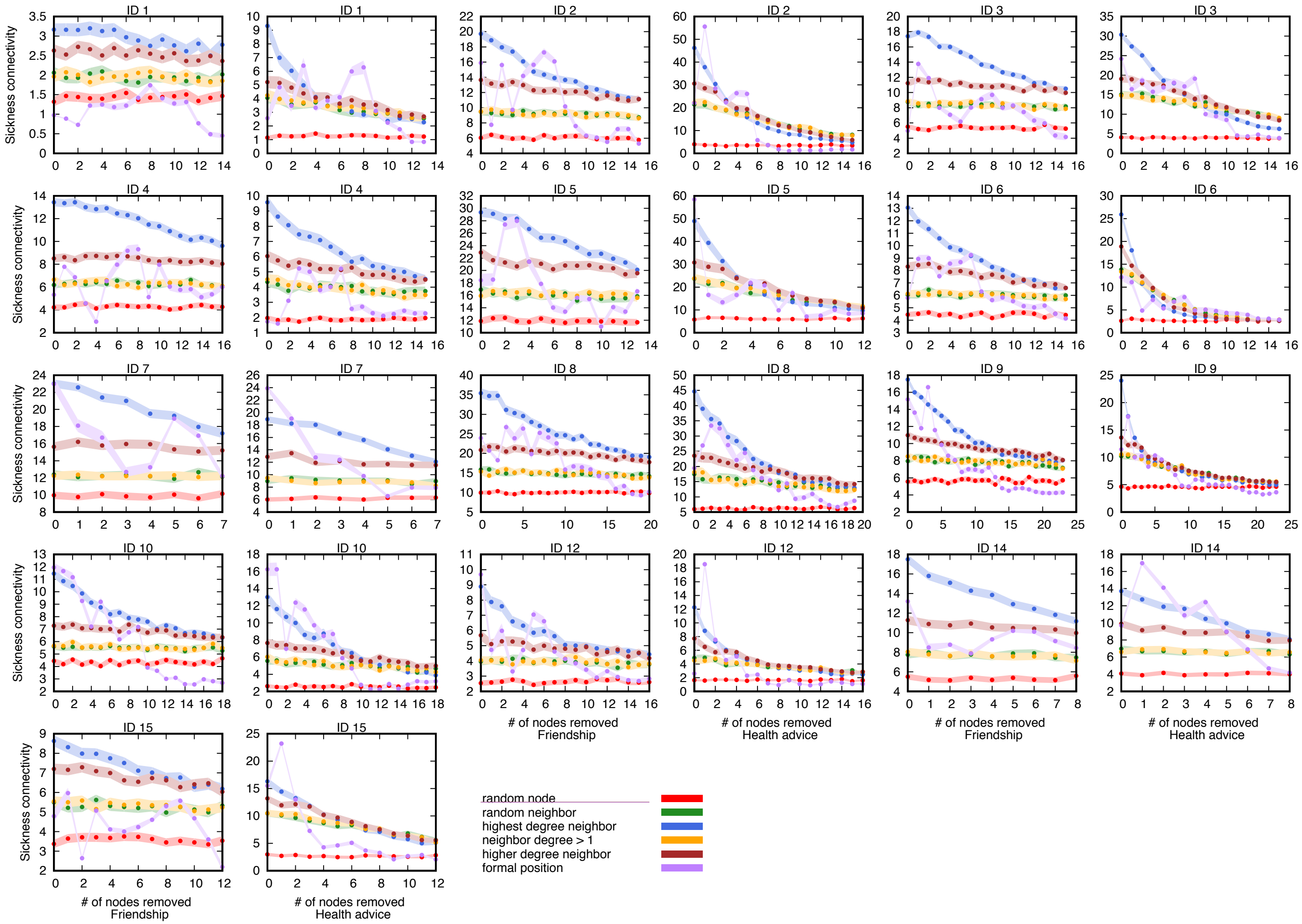


Figure S6: Avg. connectivity to sick people for 13 friendship and health advice networks. Thirteen villages are shown that were not presented in the main text. IDs correspond to project-assigned village IDs. The type of network is labeled accordingly. Sickness connectivity was defined as follows. The number of people in the neighbourhood of a node who reported diarrhea within the three months preceding the sociometric survey was divided by the degree of the node of interest. The average sickness connectivity for each node removed is shown up to the number of formal positions. One thousand iterations were run and line widths represent 95% confidence intervals.

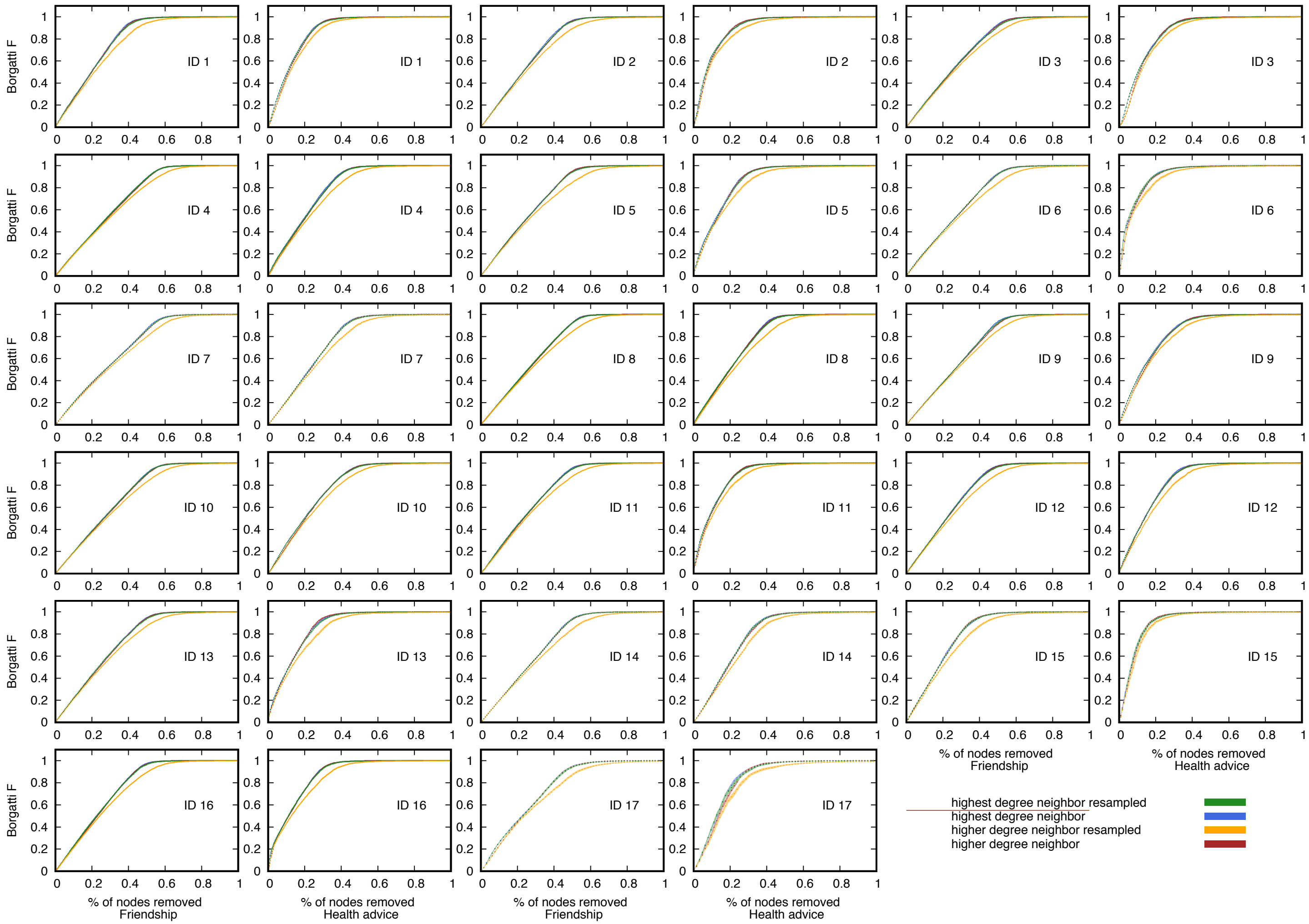


Figure S7: Acquaintance-degree strategy with node replacement. IDs correspond to project-assigned village IDs. The type of network is labeled accordingly. Line widths are greater than the 95% confidence intervals. The acquaintance-degree strategies from main text Figures 2-3 are shown here. In addition, these strategies (green and yellow) were run with one change. If a neighbour was not found, i.e. the node was an isolate, then the node was not removed from the network for the resampled highest degree neighbour strategy. This change made no difference in fragmentation efficiency since Borgatti F accounts for network fragment size. However, degree cutoffs were relaxed for the resampled higher degree neighbour strategy. If a neighbour of higher degree than the initially randomly selected node was not found then the initial node remained in the network and another node was selected until the criteria of having higher degree was met. In this case, the resampled higher degree neighbour strategy performed worse, requiring a greater percentage of nodes to induce fragmentation, than the original higher degree neighbour algorithm.

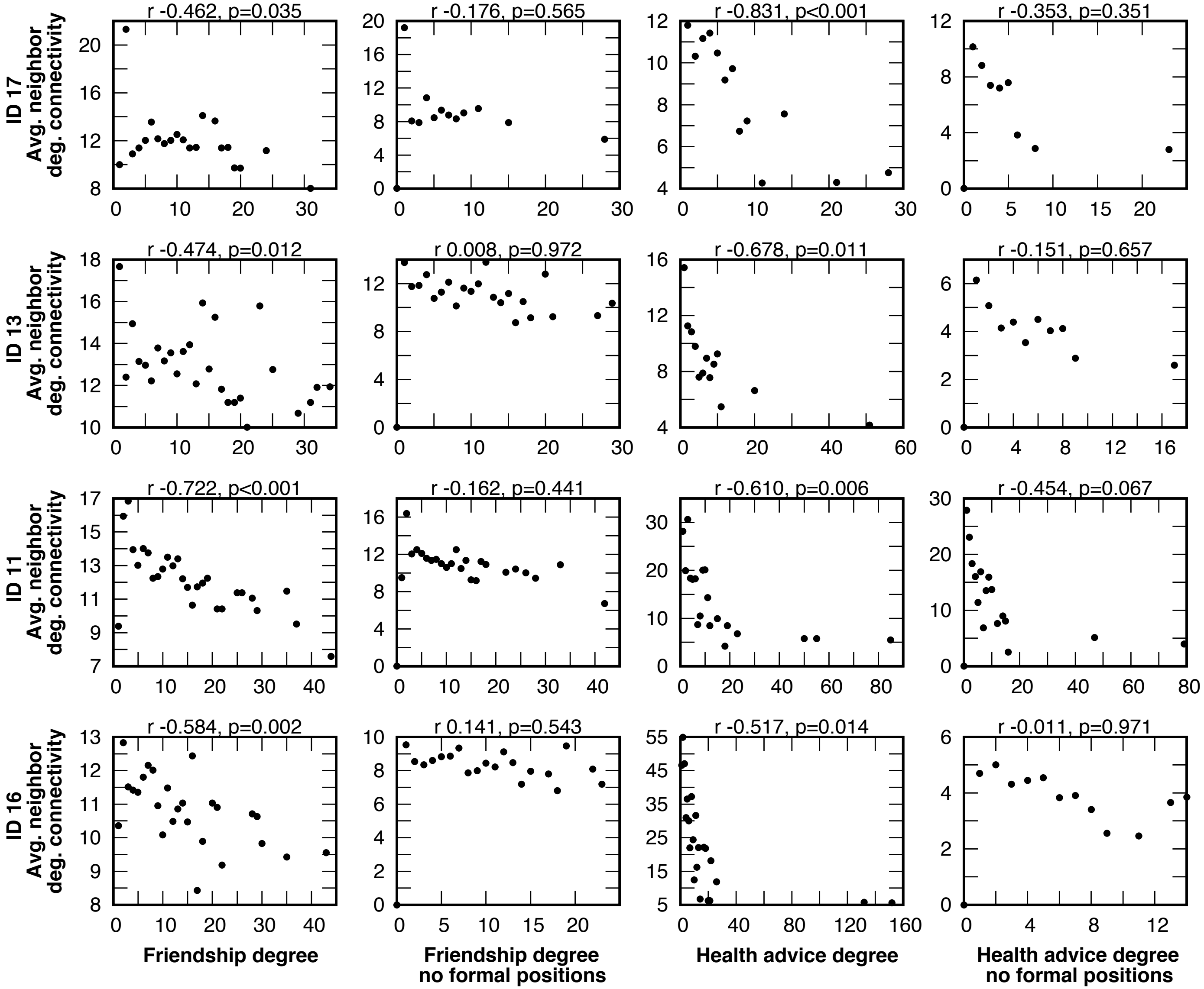


Figure S8: Degree and average neighbour connectivity correlations for 4 main text villages. Four villages are shown that had the fewest, median, 75th percentile, and greatest number of nodes. The remaining villages are shown in Figure S9. IDs correspond to project-assigned village IDs. The type of network is labeled accordingly. The Pearson correlation coefficient r of average neighbor connectivity with degree level is provided above each plot. Two plots per village are shown; one plot presents all nodes in a village and the adjacent plot shows excludes nodes with formal positions.

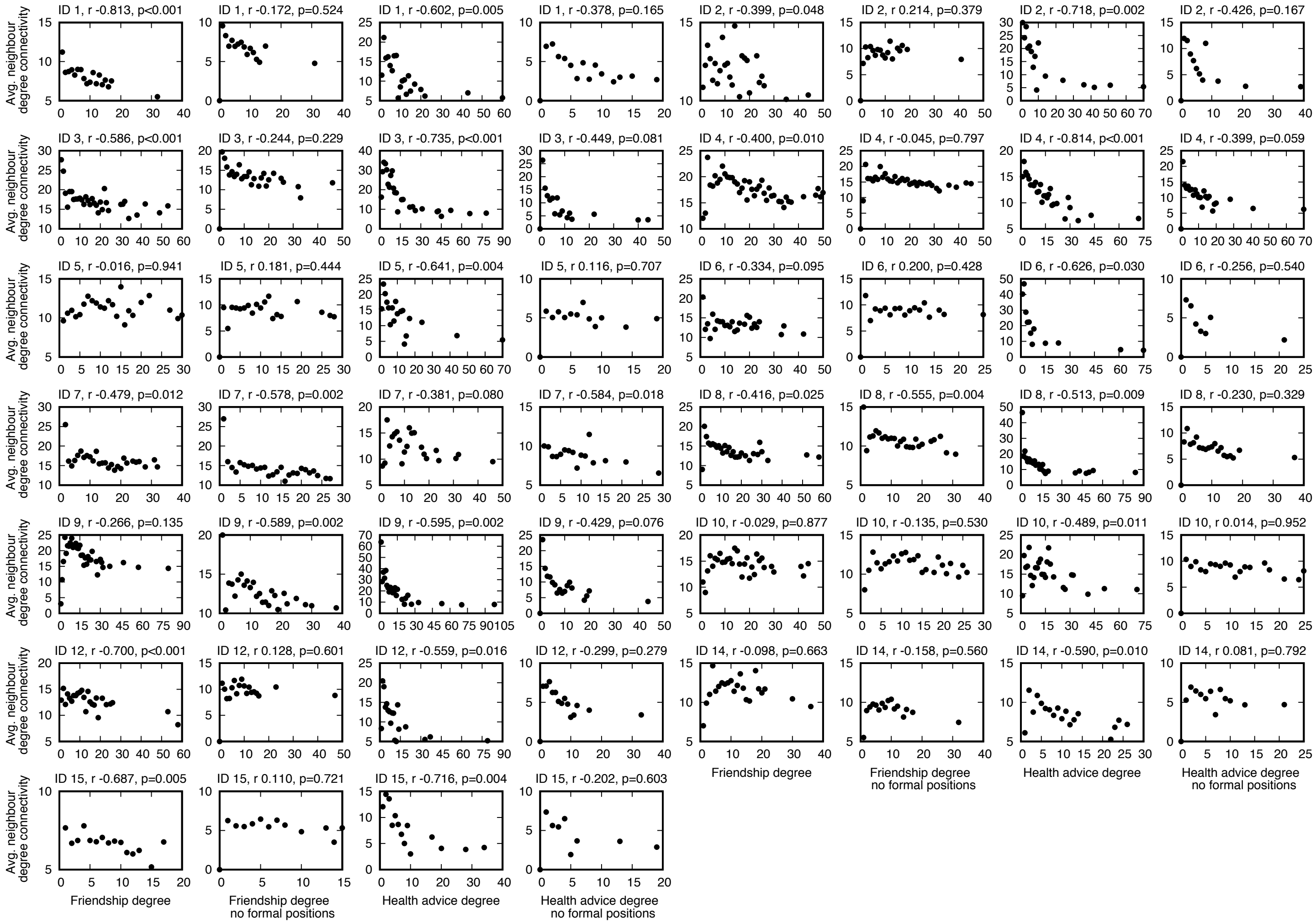


Figure S9: Degree and average neighbour connectivity correlations. Thirteen villages are shown that were not presented in the main text. IDs correspond to project-assigned village IDs. The type of network is labeled accordingly. The Pearson correlation coefficient r of average neighbour connectivity with degree level is provided above each plot. Two plots per village are shown; one plot presents all nodes in a village and the adjacent plot excludes nodes with formal positions.

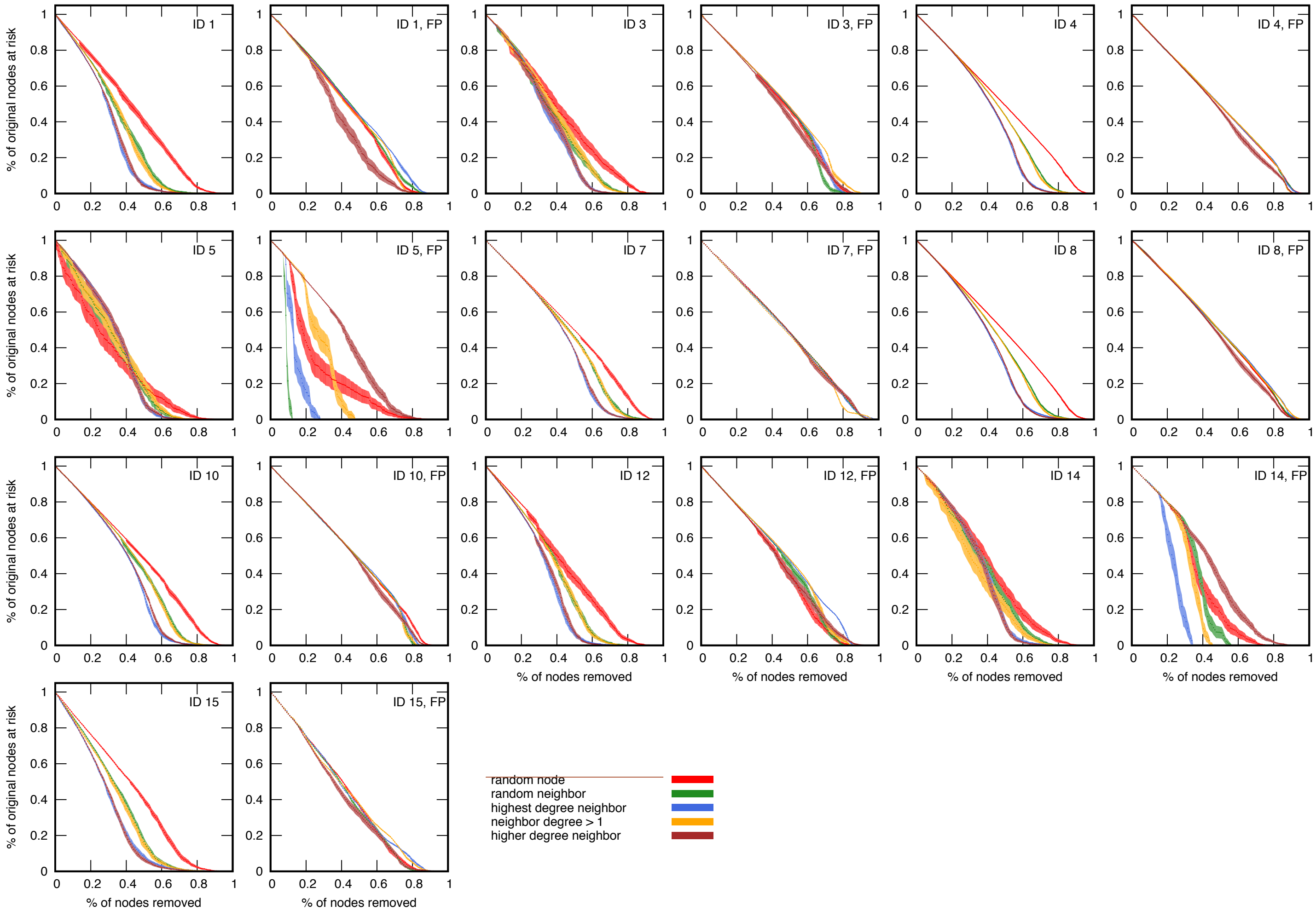


Figure S10: Health outcomes for friendship networks. Ten villages are shown that were not presented in the main text. IDs correspond to project-assigned village IDs. Three villages (IDs 2, 6 and 9) are not presented because there were zero non-complying households. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

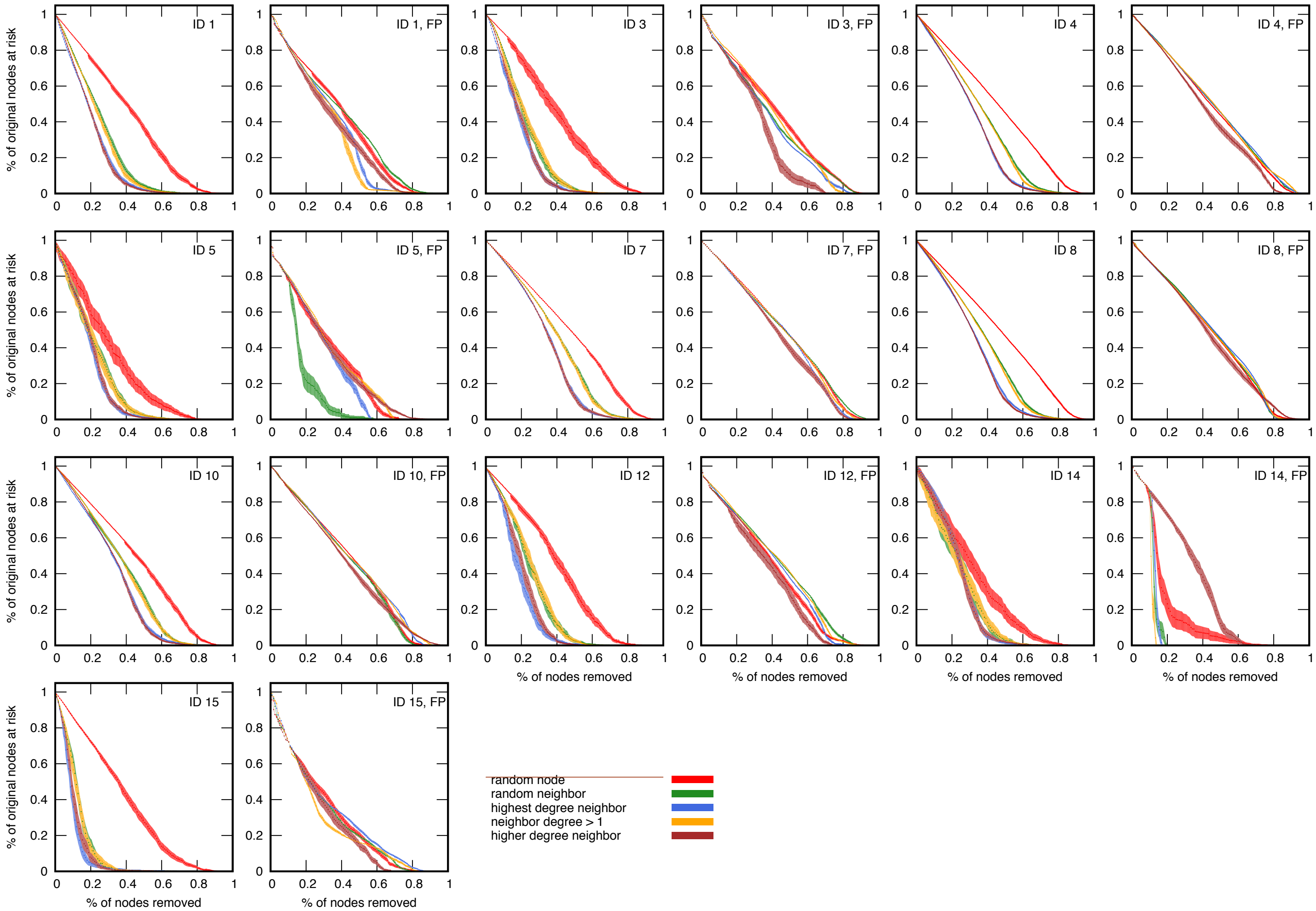


Figure S11: Health outcomes for health advice networks. Ten villages are shown that were not presented in the main text. IDs correspond to project-assigned village IDs. Three villages (IDs 2, 6 and 9) are not presented because there were zero non-complying households. If FP is noted then the formal position strategy was employed; otherwise, acquaintance and acquaintance-degree strategies were used. Line widths represent 95% confidence intervals.

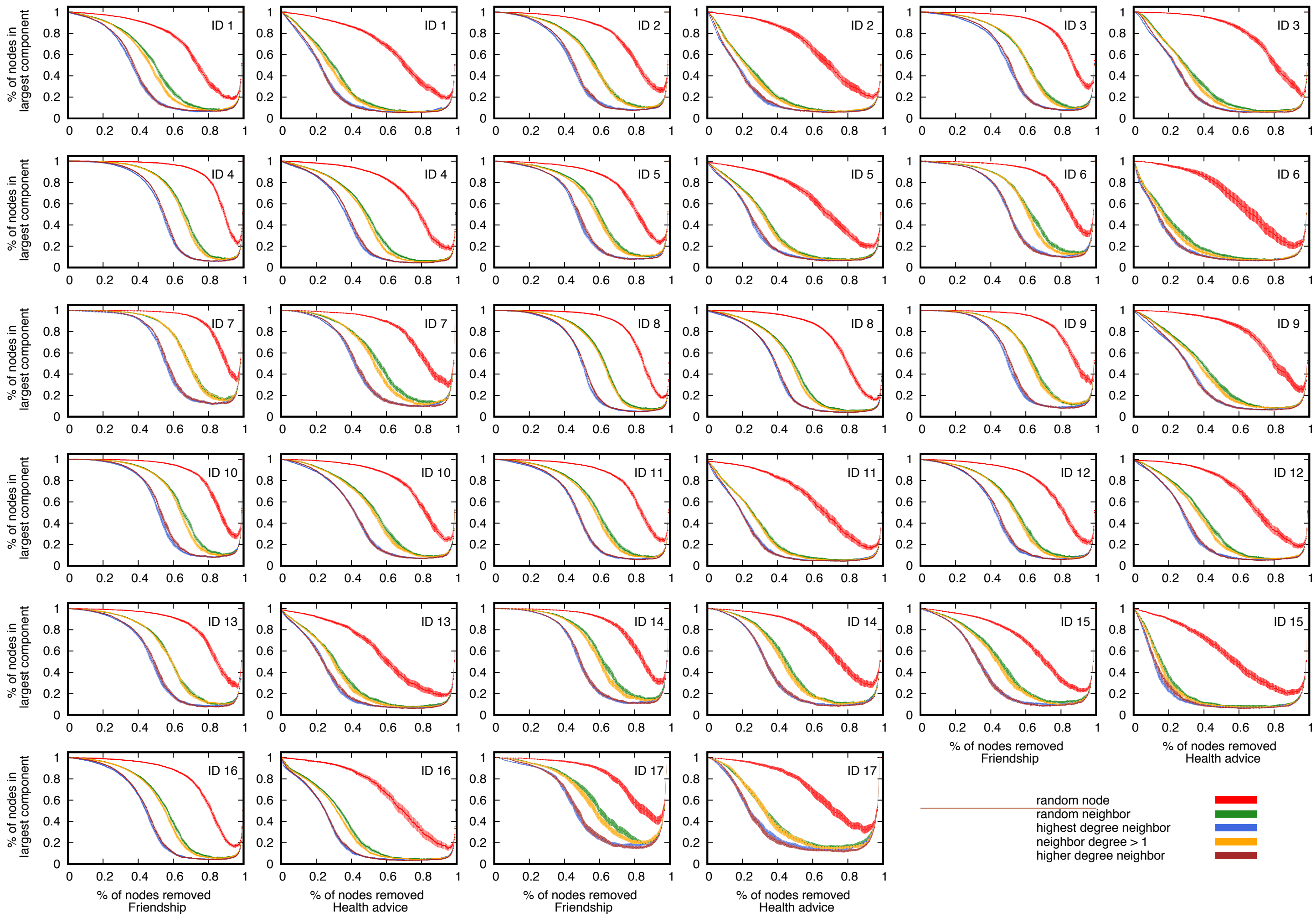


Figure S12: Percentage of nodes remaining in the largest component. IDs correspond to project-assigned village IDs. Line widths represent 95% confidence intervals.

Table S1 Households with formal positions by village and network type

	Friendship networks				*Health advice networks*			
Village ID	**All house-holds in network**	**Health workers**[a]	**Local council members**[b]	**School-teachers**	**All house-holds in network**	**Health workers**[a]	**Local council members**[b]	**School-teachers**
1	202	5	8	2	187	4	8	2
2	181	3	7	6	170	3	7	6
3	192	4	8	4	185	4	8	4
4	320	3	7	7	316	3	6	7
5	184	4	5	5	168	3	5	5
6	139	2	8	6	131	2	8	6
7	121	4	4	0	121	4	4	0
8	369	3	7	11	361	3	6	11
9	178	8	5	11	173	8	5	11
10	207	4	8	7	204	4	8	7
11	250	3	9	4	238	3	9	4
12	229	3	8	6	220	3	8	6
13	183	5	5	2	159	5	4	2
14	124	2	7	0	120	2	7	0
15	120	3	9	1	117	3	9	1
16	372	9	7	10	349	9	7	10
17	65	3	6	3	63	3	6	3

[a] Each village had two community medicine distributors, who were responsible for distributing treatment in mass drug administration. Additional households included individuals with an income-earning occupation as a health worker.
[b] Households with at least one current member of the village government.

Only households in each network (no isolates) are presented. Villages with many schoolteachers (IDs 8-9 & 16) had a private or government primary school located within the village.

Table S2 Two-sample t-tests of degree by formal position

		Friendship networks					Health advice networks				
Village ID	**Formal position**	**Obs.**	**Avg. degree**	**Std. err.**	**P-value**	**Min, Max degree**	**Obs.**	**Avg. degree**	**Std. err.**	**P-value**	**Min, Max degree**
1	No	187	5.128	0.266			173	3.809	0.232		
	Yes	15	9.000	1.447	<0.001	1 17	14	15.071	4.575	<0.001	1 61
2	No	165	7.030	0.390			154	3.818	0.324		
	Yes	16	13.500	2.449	<0.001	4 35	16	13.000	5.287	<0.001	1 70
3	No	176	10.256	0.532			169	5.497	0.443		
	Yes	16	19.688	3.355	<0.001	2 49	16	23.063	6.076	<0.001	3 81
4	No	303	12.386	0.436			300	6.357	0.347		
	Yes	17	23.471	3.763	<0.001	5 50	16	12.938	2.459	<0.001	2 35
5	No	170	7.288	0.369			155	3.826	0.247		
	Yes	14	14.357	2.180	<0.001	2 30	13	15.308	5.598	<0.001	3 71
6	No	123	8.244	0.430			115	3.348	0.243		
	Yes	16	15.625	2.666	<0.001	6 42	16	12.063	5.602	<0.001	1 76
7	No	113	12.195	0.569			113	7.265	0.431		
	Yes	8	20.250	3.411	0.001	4 31	8	20.625	5.305	<0.001	3 49
8	No	348	9.011	0.276			341	6.152	0.212		
	Yes	21	17.619	3.715	<0.001	3 58	20	20.300	5.079	<0.001	3 85
9	No	154	11.039	0.560			149	6.067	0.448		
	Yes	24	18.333	3.635	<0.001	5 80	24	14.917	4.816	<0.001	1 100
10	No	188	10.053	0.410			185	7.346	0.371		
	Yes	19	15.474	3.012	0.001	2 44	19	16.474	4.654	<0.001	1 74
11	No	234	8.175	0.373			222	4.104	0.456		
	Yes	16	13.313	1.932	0.001	5 37	16	9.688	3.385	0.004	1 55
12	No	212	7.189	0.369			203	4.394	0.240		
	Yes	17	12.941	3.400	<0.001	1 58	17	11.529	4.778	<0.001	1 82
13	No	171	8.567	0.426			148	3.257	0.203		
	Yes	12	12.750	2.903	0.02	1 32	11	9.091	4.318	<0.001	1 51
14	No	115	7.843	0.454			111	4.901	0.306		
	Yes	9	16.444	2.231	<0.001	6 30	9	14.444	2.858	<0.001	6 27
15	No	107	4.271	0.290			104	2.779	0.271		
	Yes	13	7.000	1.038	0.003	1 15	13	8.385	2.999	<0.001	1 35
16	No	346	6.879	0.226			323	3.895	0.136		
	Yes	26	14.000	2.168	<0.001	3 43	26	18.462	7.374	<0.001	1 154
17	No	53	7.226	0.732			51	4.020	0.584		
	Yes	12	10.917	1.751	0.039	4 20	12	7.917	1.520	0.008	2 21

Table S3 Physical proximity of formal position households compared to all other study households

Village ID	Households without formal positions			Households with formal positions			Two-sample t-test
	Obs.	Avg. haversine distance	std. dev.	Obs.	Avg. haversine distance	std. dev.	p-value
1	163	515.355	180.362	12	529.446	200.170	0.796
2	144	380.467	93.025	13	352.010	97.823	0.294
3	132	356.263	171.640	14	284.678	43.314	0.123
4	182	326.730	99.998	14	404.373	170.759	0.009
5	152	402.551	258.231	13	454.910	306.306	0.490
6	118	509.825	116.780	14	521.274	130.106	0.732
7	104	504.707	149.064	8	554.097	189.096	0.378
8	223	173.203	44.824	17	166.781	45.438	0.570
9	141	642.594	127.276	21	620.337	102.711	0.446
10	121	427.990	130.726	16	370.040	87.781	0.088
11	199	407.881	119.164	16	478.496	183.147	0.031
12	160	476.924	148.940	15	440.300	107.558	0.354
13	146	347.415	115.323	11	427.564	231.407	0.044
14	107	374.883	62.249	9	367.295	101.546	0.740
15	105	616.907	205.948	10	554.271	258.305	0.371
16	237	193.050	35.893	23	176.385	33.067	0.033
17	49	132.012	57.812	12	139.980	71.118	0.684

Amongst all households, 77.94% (2721/3491) had GPS waypoint data available that was matched to the household surveys. GPS waypoints were collected in November 2014. For households with individuals who had formal positions, 12.18% (33/271) did not have GPS waypoint data. The haversine distance in meters ('as the crow flies' distance) was measured between each household and every other household within the village, including those households not necessarily matched to the questionnaires. In Python v2.7, physical proximity was calculated as the average haversine distance of the household of interest to every other home in the village. Formal position households only had significantly closer physical proximity (p-value<0.05) when compared to all other households in one village (ID 16).

Table S4 Degree distributions of study networks compared to random networks

Village ID	Network type	Nodes	Edges	Avg. degree	Std. dev. of degree	Mean std. dev. of degree in an ER network of same size	Std. dev. of std. dev. of degree in an ER network of same size
1	health advice	190	458	4.821	6.097	2.156	0.077
2	health advice	170	419	4.929	7.836	2.175	0.088
3	health advice	185	648	7.005	9.891	2.582	0.136
4	health advice	316	1074	6.797	6.490	2.571	0.080
5	health advice	168	398	4.738	6.710	2.133	0.084
6	health advice	134	303	4.522	8.516	2.075	0.097
7	health advice	121	513	8.479	6.513	2.784	0.258
8	health advice	361	1287	7.130	7.239	2.636	0.075
9	health advice	173	645	7.457	10.425	2.655	0.158
10	health advice	205	840	8.195	7.912	2.791	0.156
11	health advice	240	566	4.717	7.504	2.141	0.060
12	health advice	221	543	4.914	6.429	2.182	0.069
13	health advice	173	345	3.988	4.449	1.962	0.064
14	health advice	120	336	5.600	4.471	2.291	0.145
15	health advice	117	201	3.436	4.666	1.810	0.073
16	health advice	350	915	5.229	10.906	2.263	0.049
17	health advice	63	151	4.794	4.412	2.070	0.196
1	friendship	203	568	5.596	4.032	2.321	0.090
2	friendship	182	707	7.769	5.931	2.712	0.160
3	friendship	192	1076	11.208	8.231	3.231	0.258
4	friendship	320	2115	13.219	8.590	3.549	0.207
5	friendship	184	728	7.913	5.481	2.737	0.163
6	friendship	139	640	9.209	6.182	2.911	0.258
7	friendship	121	788	13.025	6.576	3.379	0.462
8	friendship	369	1781	9.653	6.748	3.058	0.115
9	friendship	178	1080	12.135	9.446	3.343	0.307
10	friendship	207	1105	10.676	6.868	3.166	0.225
11	friendship	250	1075	8.600	5.983	2.870	0.139
12	friendship	229	885	7.729	6.680	2.721	0.129
13	friendship	183	814	8.896	6.013	2.893	0.194
14	friendship	124	538	8.677	5.490	2.817	0.261
15	friendship	120	279	4.650	3.291	2.096	0.111
16	friendship	372	1379	7.414	5.301	2.688	0.077
17	friendship	65	259	7.969	5.547	2.601	0.380

Exact numerical calculations were performed. The standard deviation of the degree in each real-world network is comparable in size to the average degree, and in some cases even larger than it. Such large standard deviations are indicative of heavy-tailed degree distributions in our study networks. The Erdős–Rényi random (ER) networks were calculated with the same number of nodes and edges as the real-world study networks. In the ER networks, the average degree is the same because we are fixing the number of nodes and edges, however the standard deviation is much smaller. The differences between the standard deviations of degree for the study networks and that of the ER networks is much larger than the fluctuations that one may expect from the sampling that gives rise to the ER networks.

Table S5 Average core numbers of nodes with formal positions

	Friendship networks				*Health advice networks*				
Village ID	**Obs.**	**Avg. core number for nodes with formal positions**	**Std. dev.**	**Max core number for all nodes**	**Obs.**	**Avg. core number for nodes with formal positions**	**Std. dev.**	**Max core number for all nodes**	**P-value from paired t-test of avg. core number of nodes with formal position**
1	15	3.467	1.060	4	14	3.357	1.082	4	0.752
2	16	5.125	0.806	6	16	2.688	1.138	4	<0.001
3	16	7.188	1.559	8	16	4.438	0.892	5	<0.001
4	17	8.118	1.409	9	16	4.375	0.885	5	<0.001
5	14	4.786	0.802	5	13	3.462	0.519	4	<0.001
6	16	5.688	0.479	6	16	2.750	0.577	3	<0.001
7	8	7.500	1.414	8	8	4.625	0.744	5	<0.001
8	21	5.619	0.740	6	20	4.350	0.671	5	<0.001
9	24	7.458	0.932	8	24	4.125	1.329	5	<0.001
10	19	6.158	1.385	7	19	4.789	2.149	7	0.001
11	16	5.750	0.577	6	16	2.813	0.834	4	<0.001
12	17	4.706	1.611	6	17	2.706	0.588	3	<0.001
13	12	5.000	1.758	6	11	2.636	1.027	4	<0.001
14	9	5.778	0.441	6	9	3.889	0.333	4	<0.001
15	13	3.231	0.832	4	13	2.231	0.725	3	0.004
16	26	4.731	0.604	5	26	3.192	0.939	4	<0.001
17	12	5.083	1.084	6	12	3.583	0.793	4	<0.001

Table S6 Two-sample t-tests of degree by noncomplying household

Village ID	Noncomplying household[a]	*Friendship networks* Obs.	Avg. degree	Std. err.	P-value	*Health advice networks* Obs.	Avg. degree	Std. err.	P-value
1	No	198	5.338	0.277		183	4.596	0.457	
	Yes	4	9.250	1.797	0.0484	4	7.250	2.175	0.3951
2	No	181				170			
	Yes	0				0			
3	No	190	11.011	0.594		183	7.016	0.754	
	Yes	2	14.000	2.000	0.6075	2	7.000	2.000	0.9982
4	No	302	12.772	0.489		298	6.641	0.377	
	Yes	18	16.389	2.044	0.0806	18	7.500	1.023	0.5812
5	No	183	7.842	0.403		167	4.713	0.535	
	Yes	1	5.000			1	5.000		
6	No	139				131			
	Yes	0				0			
7	No	113	12.894	0.615		113	8.274	0.638	
	Yes	8	10.375	2.652	0.299	8	6.375	1.133	0.4351
8	No	344	9.363	0.338		336	6.976	0.382	
	Yes	25	11.400	2.208	0.1421	25	6.400	1.990	0.7012
9	No	178				173			
	Yes	0				0			
10	No	201	10.587	0.482		198	8.303	0.587	
	Yes	6	9.333	2.333	0.6573	6	4.667	0.882	0.2836
11	No	243	8.564	0.384		231	4.498	0.502	
	Yes	7	6.429	1.986	0.3516	7	3.857	0.553	0.8248
12	No	226	7.650	0.437		217	4.982	0.446	
	Yes	3	5.000	1.155	0.4867	3	2.333	0.333	0.4872
13	No	172	8.843	0.454		149	3.738	0.385	
	Yes	11	8.818	2.173	0.9895	10	2.500	0.428	0.4081
14	No	122	8.525	0.496		120	5.622	0.421	
	Yes	2	5.000	4.000	0.3684	1	5.000		
15	No	114	4.561	0.301		111	3.486	0.455	
	Yes	6	4.667	1.202	0.9376	6	1.833	0.307	0.4026
16	No	345	7.339	0.289		324	5.046	0.636	
	Yes	27	7.852	0.789	0.6283	25	4.120	0.343	0.6865
17	No	57	7.596	0.718		55	4.636	0.653	
	Yes	8	10.125	2.416	0.2356	8	5.625	0.962	0.5759

[a] In total, there were 129 noncomplying households. Three villages (IDs 2, 6, & 9) did not have any noncompliance attributable to adverse drug effects. In the other 14 villages, noncompliance widely varied (Avg. 9.214, std. dev. 8.541). Village IDs 5 & 14 only had one noncomplying household.

References

1 Chami, G. F. *et al.* (2016) Profiling nonrecipients of mass drug administration for schistosomiasis and hookworm infections: a comprehensive analysis of praziquantel and albendazole coverage in community-directed treatment in Uganda. *Clin Infect Dis* 62(2): 200-207

2 Cohen, R., Havlin, S. & ben-Avraham, D. (2003) Efficient immunization strategies for computer networks and populations. *Phys Rev Lett* 91(24): 247901

3 Gallos, L. K., Liljeros, F., Argyrakis, P., Bunde, A. & Havlin, S. (2007) Improving immunization strategies. *Phys Rev E* 75(4): 045104

4 Holme, P. (2004) Efficient local strategies for vaccination and network attack. *EPL* 68(6): 908

5 Fenwick, A. *et al.* (2009) The Schistosomiasis Control Initiative (SCI): rationale, development and implementation from 2002-2008. *Parasitology* 136(13): 1719-1730

6 Freeman, L. C. (1979) Centrality in social networks conceptual clarification. *Soc Net* 1(3): 215-239

7 Albert, R., Jeong, H. & Barabási, A.-L. (2000) Error and attack tolerance of complex networks. *Nature* 406(6794): 378-382

8 Girvan, M. & Newman, M. E. J. (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99(12): 7821-7826

9 Holme, P., Kim, B. J., Yoon, C. N. & Han, S. K. (2002) Attack vulnerability of complex networks. *Phys Rev E* 65(5): 056109

10 Borgatti, S. P. (2006) Identifying sets of key players in a social network. *Comput Math Organ Theory* 12(1): 21-34,

11 Chen, Y. *et al.* (2007) Percolation theory applied to measures of fragmentation in social networks. *Phys Rev E* 75(4): 046107

12 Schneider, C. M., Moreira, A. A., Andrade, J. S., Havlin, S. & Herrmann, H. J. (2011) Mitigation of malicious attacks on networks. *Proc Natl Acad Sci USA* 108(10): 3838-3841

13 Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y. & Shir, E. (2007) A model of Internet topology using k-shell decomposition. *Proc Natl Acad Sci USA* 104(27):11150-11154

14 Webster, J. P., Molyneux, D. H., Hotez, P. J. & Fenwick, A. (2014) The contribution of mass drug administration to global health: past, present and future. *Philos Trans R Soc Lond, B, Biol Sci* 369(1645)

15 Krentel, A., Fischer, P. U. & Weil, G. J. (2013) A review of factors that influence individual compliance with mass drug administration for elimination of lymphatic filariasis. *PLoS Negl Trop Dis* 7(11): e2447

16 Kelly-Hope, L. A., Thomas, B. C., Bockarie, M. J. & Molyneux, D. H. (2011) Lymphatic filariasis in the Democratic Republic of Congo; micro-stratification overlap mapping (MOM) as a prerequisite for control and surveillance. *Parasit Vectors* 4(1):178

17 Parker, M. & Allen, T. (2011) Does mass drug administration for the integrated treatment of neglected tropical diseases really work? Assessing evidence for the control of schistosomiasis and soil-transmitted helminths in Uganda. *Health Res Policy Syst* 9(3): 3
18 Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2013) The Diffusion of microfinance. *Science* 341(6144)
19 Valente, T. W. & Davis, R. L. (1999) Accelerating the Diffusion of Innovations Using Opinion Leaders. *Ann Am Acad Pol Soc Sci* 566(1): 55-67
20 Rogers, E. M. (2003) *Diffusion of innovations.* 5 edn, (Free Press).